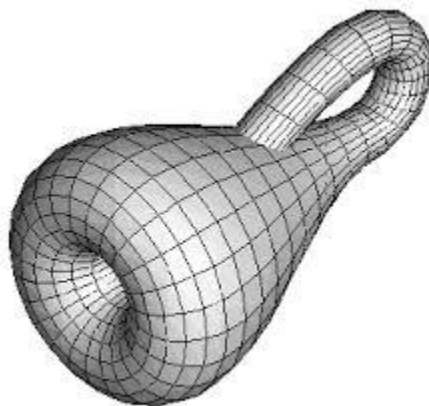# Continuum

The St. Clair County Community College Journal of Mathematics

Continuum
The St. Clair County Community College Journal of Mathematics

Articles

Dedicated to the memories of James Mattiello, PhD, and Diane Bedard

Welcome to the second issue of Continuum. Our previous issue, dated Winter 2016, is available through the St. Clair County Community College Library. We're proud to bring you the best original work in mathematics, both pure and applied, from the faculty (both full and part time) at St. Clair County Community College. We're very excited for this edition to also feature the work of one of our outstanding students, Rebekah Muzzi.

We hope to bring out a new edition on an annual basis. Last time around, we were pleased to be able to say that all seven of our full-time mathematics professors contributed. This time, our focus was including contributors from outside our discipline. The results: two applied mathematics articles, one from Dr. Janice Fritz in our biology discipline, and another from Dr. James Mattiello, an adjunct professor who passed away just after completing this remarkable article. We are particularly proud to present this work by a truly brilliant physicist.

The articles here are intended mostly for mathematics enthusiasts but also offer teaching ideas, and the background knowledge required varies. Even if you know but little mathematics, the article about the Online Encyclopedia of Integer Sequences should be approachable. Rebekah's article, on the other hand, is highly recommended for readers with a fair amount of calculus knowledge. We've also included a couple of brief mathematician biographies. In future issues we would love to include responses to any inquiries, or articles tailored to reader requests. One idea we may incorporate is an "Ask a Mathematician" column, in which we answer your questions.

If you would like to comment, ask us anything, or offer any input whatsoever, we would love to hear from you. Please email or write to the addresses below. If you wish your comments forwarded to any of our authors, we will do that. Please include explicit permission to publish your letters, if you wish to grant it, and we may print some letters in the next edition.

Thank you for your interest, and we hope you find something to stimulate your imagination and make you think.

Paul Bedard: pbedard@sc4.edu
Nick Goins: ngoins@sc4.edu

Or write us care of:

St Clair County Community College
323 Erie Street
PO Box 5015
Port Huron, MI 48061-5015

Fun with *The Online Encyclopedia of Integer Sequences* (OEIS)

Paul Bedard

Introduction

The first mathematics a child learns is an integer sequence.

$$1, 2, 3, 4, 5, \cdots$$

It's just counting, or "saying your numbers." Later, students learn their "times tables" – merely more sequences:

$$4, 8, 12, 16, 20, \cdots$$

This is the list of the multiples of 4. An old Persian folktale tells of a foolish king who offered a wise man any reward he could name. The wise man asked for a chess board with one grain of wheat on the first square, two on the second, four on the third, and so on – each square doubling the number of grains. The king agreed, not knowing how quickly exponential functions grow. The sequence

$$1, 2, 4, 8, 16, \cdots$$

grows to 2^{63} or 9,223,372,036,854,775,808 by the 64th entry (the last square on the chess board.)

There is a fascination in lists, and lists with a pattern are more intriguing still. In elementary school, my friend Bill Mullan and I used to challenge each other daily. One of us would call out the beginning of a sequence across the playground, and the other, wherever he was, would have to continue it if he could. Neither of us ever stumped the other. There are so many possibilities! One of my favorites is the "pancake sequence."

Question. What is the largest number of pieces a pancake can be cut into with one straight cut?

Of course, there is no way to get anything other than two pieces. What about two cuts? Draw a circle and then draw straight lines across the circle. You will see that you can get 4 pieces with two cuts, if the cuts intersect each other. What about three cuts? If you cut the pancake like a pizza, you will get 6 pieces. But, if the first two cuts intersect each other and the third cut does NOT pass through the point of intersection of the first two – voila! Seven pieces! So,

$$2, 4, 7, \cdots$$

is the beginning of the pancake sequence.

The First Four Terms Of The Pancake Sequence
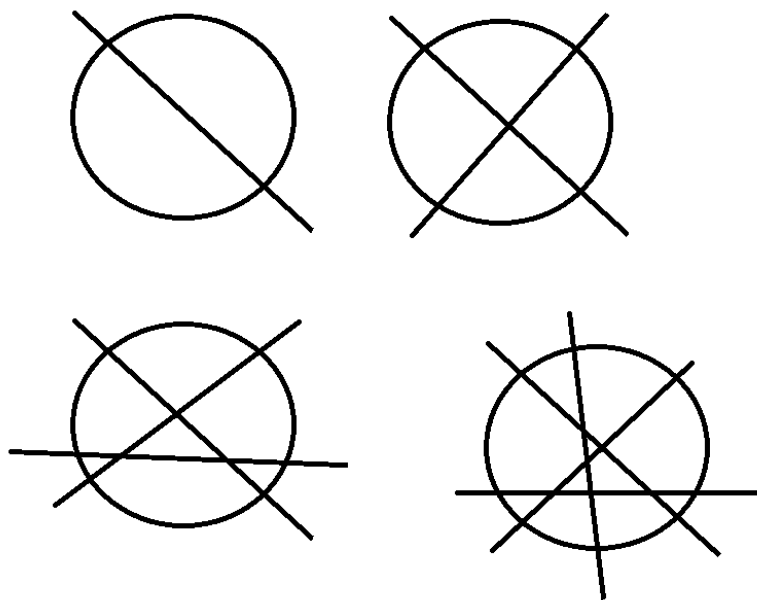
Exercise. Find the next term in the pancake sequence.

Here is a very challenging sequence that will require some "outside the box" thinking!

$$3, 3, 5, 4, 4, 3, 5, 5, 4, 3, 6, \cdots$$

I wish I had thought of this one for Bill Mullan! But it isn't really fair. You are probably looking for a mathematical pattern. I have simply listed the number of letters in the English words

$$one, two, three, four, five, six, seven, eight, nine, ten, eleven, \cdots$$

By now you must be wondering about the "…" I keep using. These three dots are called an **ellipsis**, and they represent the fact that the sequence continues, following the same pattern, forever. The ellipsis can sometimes be misleading, so it should never be used until a clear pattern has been established. For instance, if I write 2,4,6,…, then the reader spots quickly that I am listing even integers, and assumes that the sequence continues …,8, 10,12, … However, I may have intended some other pattern that just coincidentally starts with 2, 4, 6.

Perhaps you are curious about the fourth entry in the pancake sequence. It's time I introduce the real star of this story – the OEIS, the Online Encyclopedia of Integer Sequences.

https://oeis.org/

With this wonderful tool, we can explore all sorts of exciting sequences. What's more, the Encyclopedia explains how the numbers are found, and also mentions other ways the same list can be created. For instance,

$$1, 3, 6, 10, 15, \cdots$$

is a sequence called the "**triangular numbers**." Think of an array of bowling pins:

```
   o      o       o       o
      o        o        o
        o        o
           o
```

There are ten pins. But what if you made an array with only three rows of pins?

```
   o      o       o
      o        o
        o
```

But the triangular numbers are also:

- The number of legal ways to insert a pair of parentheses in a string of *n* letters.
- The number of tiles in a set of double – *n* dominoes.
- The number of ways a chain of *n* non-identical links can be broken up.
- The maximum number of intersections of n+1 lines which may only have 2 lines per intersection point.
- Number of ways two different numbers can be selected from the set {0,1,2,...,n} without repetition, or, number of ways two different numbers can be selected from the set {1,2,...,n} with repetition.

If any of the above doesn't make sense, look it up! (The above bulleted points are not my words: they are a direct quote from the OEIS.)

The OEIS was started by Neil Sloane in 1964. "The sequence database was begun by Neil J. A. Sloane in early 1964 when he was a graduate student at Cornell University in Ithaca, NY." (OEIS). Sloane began to curate interesting sequences by writing them on file cards. 2,372 of these sequences were recorded in a book, A Handbook of Integer Sequences", Academic Press, NY in 1973. Now, anyone who discovers a sequence can propose it for addition to the database. (It is not, however, easy to come up with one nowadays that no one else has thought of! More on this later.) As of this writing, the OEIS contains Read more about the history at https://oeis.org/wiki/Welcome.

Exercise 1: Derive the next 5 terms in the sequence of triangular numbers.

Exercise 2: Derive the 4th term in the pancake sequence.

Exercise 3: Find each of the sequences given above in the OEIS and state their catalog reference.

Some Basic Sequence Facts

Before we start exploring, I will introduce some notation and terminology. A **sequence** is a lsit of numbers, separated by commas. A particular number in a sequence is called an **entry** of the sequence. An **integer** is a number which is either a counting number (1,2,3,4, …), or the opposite (negative) of a counting number, or zero. An **integer sequence** is a sequence all the entries of which are integers.

A general entry in a sequence (that is, an entry we want to discuss without necessarily knowing its value) is usually referred to by the lower case letter a, with a subscript denoting the placement within the sequence. So, a_5 is the fifth entry in a sequence.

The expression $\{a_n\}$ represents an entire sequence.

A formula for a sequence can be given by referring to the subscript n as though it were a variable For example, the sequence of odd positive integers

$$1, 3, 5, 7, 9, 11, \cdots$$

can be generated using the formula $a_n = 2n - 1$. To find any entry, say a_5, we simply evaluate the formula for that value of n.

$$a_5 = 2(5) - 1 = 9$$

There is another way a formula for a sequence can be given: a **recursive formula** shows how to obtain an entry in a sequence by using previous entries. The most famous recursive formula is the one that defines the **Fibonacci sequence**:

$$a_n = a_{n-1} + a_{n-2}$$

In ordinary words, this formula says that each entry in the Fibonacci sequence is the sum of the two entries that come right before it. But this implies that we have to be told what the very first two entries are, since there is no way to use the recursive formula to obtain them The first two entries are both 1's.

Why Just Integers?

Many important and fascinating sequences have noninteger entries. For instance, the **harmonic sequence**

$$\frac{1}{1}, \quad \frac{1}{2}, \quad \frac{1}{3}, \quad \frac{1}{4}, \quad \frac{1}{5}, \quad \frac{1}{6}, \cdots$$

Such sequences are well worth our consideration! However, there are two important reasons to look at integer sequences separately. Firstly, it will be far easier to get a hold of some of the concepts we need if we restrict ourselves to just the integers. Secondly, many important facts about noninteger sequences can be discovered by looking at related integer sequences. Let's make this our first exploration.

Let's take a sequence that is decidedly nonintegral: the sequence of square roots of positive integers. This sequence is defined by the formula $a_n = \sqrt{n}$. The first few terms are

$$\sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}, \sqrt{5}, \cdots$$

Some of these entries are in fact integers. The noninteger entries are all irrational numbers (this is moderately easy to prove!) An irrational number has a decimal representation that does not terminate or repeat. In other words, we cannot write out all the digits of an irrational number.

Let's write out the first four digits of the numbers in our sequence:

$$1.000, 1.414, 1.732, 2.000, 2.236, \cdots$$

How can we make this an interesting sequence? Before I explain this, please do the following: using a computer or calculator, list the first 20 square roots, with only three decimal places after the decimal point for each. I will leave some space for your sequence.

Okay, now we will turn this into an integer sequence by listing only the first digit of each square root. In other words, we will take the "integer part" of each of these numbers.

$$1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, \ldots$$

This is sequence **A000196** in the OEIS. What useful information can this list provide us with? Looking it up, we see that it also tells us the number of perfect square less than or equal to n. In other words, since $a_8 = 2$, there are two perfect squares less than 8 (they are 1 and 4, of course.)
Let's prove this.

Theorem 1: The integer part of the square root of a positive integer is equal to the number of perfect square less than or equal to that integer.

Proof. Let n be a positive integer. Suppose there are k perfect squares less than or equal to n. Then these perfect squares are

$$1^2, 2^2, 3^2, \cdots, k^2$$

If $k^2 = n$ then $k = \sqrt{n}$ and the integer part of $\sqrt{n}$ is k. This would satisfy the theorem. So let's assume $k^2 < n$.
Let us say that the integer part of $\sqrt{n}$ is m
Suppose $m > k$.
Then $m^2 > k^2$ but we know that $m^2 < n$. Since $k^2 < n$ we know that $k^2 < m^2 < n$. Then m^2 would be a perfect square between k^2 and n, which is a contradiction. So $m \leq k$.
Suppose $m < k$. Then $m^2 < k^2$. Since m is the integer part of $\sqrt{n}$, $(m + 1)^2 > n$. Since there are no perfect squares between m^2 and $(m + 1)^2$ and $m^2 < n < (m + 1)^2$, then there are k perfect squares less than $(m + 1)^2$. But m^2 is one of these, and the largest of these is k^2, so $m^2 \leq k^2$. This is a contradiction.
Therefore $m = k$.

Let's go back to the harmonic sequence for a moment. What sequence of integers naturally arises from this? Consider the **partial sums** of $a_n = 1/n$.

$$\frac{1}{1} = 1$$

$$\frac{1}{1} + \frac{1}{2} = \frac{3}{2}$$

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$$

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{50}{24} = \frac{25}{12}$$

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} = \frac{137}{60}$$

Let's consider the sequence of just the numerators of these fractions:

$$1, 3, 11, 25, 137, \cdots$$

This turns out to be sequence **A001008** in the OEIS. We learn that for prime numbers p greater than 3, p^2 evenly divides $a(p-1)$ for any a in this sequence! We won't prove that here. But it is interesting that a number theoretical fact about primes emerged from our investigation into this sequence.

Exercise 4: How many perfect squares are less than 11,208?

Finding More Sequences

I enjoy typing random strings of integers into the OEIs and seeing what comes up. It's truly amazing. Let's type some integers at random, with duplications:

$$1, 1, 1, 2, 2, 3, 4, 4, 5$$

The results?

- The initial digits of perfect cubes (starting with 10^3);

- Number of partitions of n into distinct parts, none being 7;

- Number of partitions of n into 8 primes;

- Number of partitions of n into 9 primes;

- Number of partitions of n into 8 primes;

- Expansion of

$$\frac{1}{(1-x)(1-x^2)(1-x^5)}$$

- Number of 7's in all the partitions of n into distinct parts.

This exploration has introduced us to the idea of partitions! Let's explore that a little further.

Definition. A **partition** of a positive integer is a presentation of the integer as a sum of integers.

Examples.

    1
    2 = 1 + 1
    3 = 2 + 1 = 1 + 1 + 1
    4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1

Including the integer itself, we have seen that 1 has one partition, 2, has two, 3 has three, and 4 has five. Let's add one more before we check OEIS.

$$5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1$$

So, our sequence representing "the number of partitions of n" is

$$1, 2, 3, 5, 7, \cdots$$

Let's enter it! It's **A000041.**

Now that we have learned a little about partitions, we can look at some of the other results.

Number of partitions of n into distinct parts, none being 7. **A015754**

Distinct parts means we cannot re-use integers. Going back to our examples above,

    1
    2 = 1̶ ̶+̶ ̶1̶
    3 = 2 + 1 = 1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶
    4 = 3 + 1 = 2̶ ̶+̶ ̶2̶ = 2̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ = 1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶
    5 = 4 + 1 = 3 + 2 = 3̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ = 2̶ ̶+̶ ̶2̶ ̶+̶ ̶1̶ = 2̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ = 1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶ ̶+̶ ̶1̶

Adding $6 = 5 + 1 = 4 + 2 = 3 + 2 + 1$

The sequence of partitions of n into distinct parts is therefore

$$1, 1, 2, 2, 3, 4, \cdots$$

Checking the OEIS for confirmation, this is indeed **A000009**; wow! Only the ninth sequence on the list! (Note that our sequence starts with the seventh Padovan number.)

While I was checking for this, I encountered the Padovan sequence, **A000931** – this is brand new to me! What can we learn here?

The OEIS defines the Padovan numbers recursively – just like the Fibonacci numbers! The recursive formula is

$$a(n) = a(n-2) + a(n-3), a(0) = 1, a(1) = a(2) = 0$$

Exercise: Write out the first 15 Padovan numbers. Check your work in the OEIS!

Let's look at just one of many uses of this sequence, quoted directly from the OEIS:

```
a(n) counts closed walks from a vertex of a unidirectional triangle
containing an opposing directed edge (arc) between the second and third
vertices. Equivalently the (1,1) entry of A^n where the adjacency matrix
of digraph is A=(0,1,0;0,0,1;1,1,0). - David Neil McGrath, Dec 19 2014
```

What does this mean? Well, here we have entered the exciting world of graph theory! Graph theory deals with the study of objects called **graphs**, which are collections of points (or vertices, or nodes) and lines (or edges). Draw a few points. Connect some of them (or all of them) with edges. You have created a graph! Graphs are also called networks, and they can model the flow of traffic, information, or just about anything, from point to point. Two of the most famous problems addressed by graph theory are the Four Color Problem and the Konigsberg Bridge Problem. (I suggest you look these up!) A **directed graph** or **digraph** is a graph where each edge has an associated direction. A directed edge can be called an **arc**. Directed graphs can be used to model the spread of ideas, diseases, or rumors from person to person or machine to machine in a network.

Let's try to figure out David McGrath's idea.
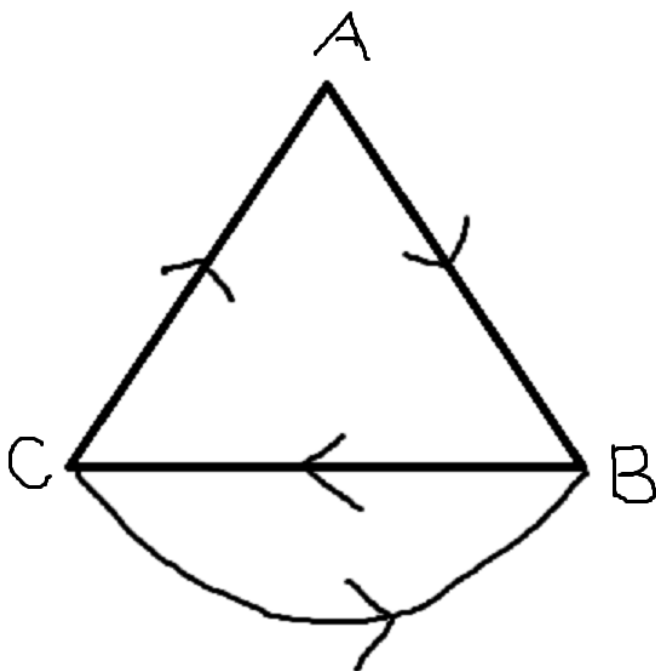
We now need some more definitions. A **walk** is a sequence of vertices that we can reach along edges. So ABC would be a walk and so would ABCAB. A **closed walk** is a walk that starts and ends at the same vertex.

Starting with A, we could return to A by the closed walk ABCA. This walk includes three edges, edge AB, edge BC, and edge AC. Therefore it has a **length** of 3.

Is there any other way to do it?

There are no closed walks of length 4.

Can you find any other walks of lengths 3 or 4?

There are no closed walks of length 4.

What about length 5?

ABCBCA is a closed walk of length 5.

Let's put this together in a chart.

n	closed walks of length n
0	1*
1	0
2	0
3	1
4	0
5	1
6	1
7	1
8	2

*(This means there is one way to get from A to A with a walk of length 0 – just stay at A!)

Exercise. What are the two closed walks of length 8?

Answer: ABCABCBCA and ABCBCABCA.

The adjacency matrix of a digraph is a matrix with an entry 1 if a path exists from the vertex labelling the row of the matrix to the vertex labelling the column, and entry 0 if no such path exists. The adjacency matrix for this digraph is

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

The 1 in the middle of the first row indicates that 1 walk of length one exists from vertex A to vertex B. The zero in the top right indicates that no walk of length one exists to get you from A to C.
For advanced readers: using matrix multiplication, the powers of the adjacency matrix give the numbers of walks of longer lengths. Thus, squaring the matrix gives the number of walks of length two, cubing gives length three walks, etc. Does McGrath's last statement now make sense?)

To convince the skeptical reader that I am truly using random strings of numbers and not carefully selecting interesting sequences, for my next example I will use the numbers of letters in the first five words of this sentence.

$$2, 8, 3, 9, 6, \cdots$$

It comes up! It is, among other things, "The spiral of Champernowne read by the South-Southwest ray." I have never before heard of this. It is also part of the decimal expansion of Soldner's constant, and part of the decimal expansion of

$$\frac{6 + 2\sqrt{10}}{3}$$

Let's try again. Let's take the counting numbers and reverse their order, in pairs.

$$2,1\ 4,3,6,5,\cdots$$

Well, OEIS comes up with our pattern, but scrolling down, I find something interesting.

A096779: Smallest number not occurring earlier having no common digits with n in its decimal representation.

Let's see what this means. The first entry ($n = 1$) of our sequence is 2. 2 is the smallest number which has no 1's as digits. The second entry ($n = 2$) is 1. One is the smallest number with no 2's as digits.

Question: Will this sequence continue to swap natural numbers in the pairwise fashion that we started with? Answer: No!

Exercise: Find the first 15 entries in this sequence.

Challenge exercise: This sequence is finite! It has a last entry. Why? And what is this last entry? See OEIS for answers.

Using Sequences to Solve Problems

As we have seen, integer sequences can help us solve many kinds of problems. Let's see if we can use the OEIS to solve some more.

Problem. How many ways are there to seat 10 people in a circle (say, at a table?)

Circular arrangements are trickier than linear ones. Let's take the advice of George Polya. In his famous book <u>How To Solve It</u>, Polya suggests the strategy of solving a simpler, related problem first.

So let's start with one, two, or three people. One person can be seated in one way. Two people can also be seated in one way – and this differs from lining people up. With a linear arrangement, AB is different from BA. But arranged around a round table, AB is indistinguishable from BA. Two arrangements around a table can be considered different if, for any one person in one arrangement, the people to her right and left are different than in the other arrangement.

What about three people? Now there are two arrangements. One is ABC, or the equivalent BCA or CAB, and the other is ACB or its equivalents CBA or BAC.

Perhaps you have spotted a pattern. If we move any person from the end of the list to the beginning, it is an equivalent circular arrangement. So, for four people,

$$ABCD = BCDA = CDAB = DABC$$

But if we start with AC we get a different arrangement:

$$ACBD = CBDA = BDAC = DACB$$

We could also start with AD. What if we start with B? We already have BCDA and BDAC. What about BA?

$$BACD = ACDB = CDBA = DBAC$$

We also need to count arrangements starting with ADC, ADB, and BCA.
WE find that there are 6 possible arrangements. (Convincing yourself of this may take a little work!)
So our sequence of answers for 1,2,3, or 4 people is

$$1, 1, 2, 6, \cdots$$

Entering this into the OEIS gives us sequence **A000142** – the **factorial** numbers. Circular arrangements represent one of many descriptions. Note that the number of circular arrangements of n objects is $(n-1)!$
Now the answer to our problem is available without our needing to perform any calculations. The number of arrangements of ten people at a round table is 362,880.

While we are here, let's note that
.

$$1, 1, 2, 6, \cdots$$

is also the beginning of an entirely different sequence. **A003418** lists the least common multiple of the set of numbers $\{1, 2, 3, \cdots, n\}$. In other words, for $n = 4$ our sequence entry is 6. 6 is the least common multiple of 1, 2, 3, 4.

Exercise. Find the first 15 terms of **A003418.**

But these four numbers 1, 1, 2, 6 are also part of another sequence, the so-called **primorial numbers.** The nth primordial number is the product of all primes less than or equal to n. For this to make sense, we need to include 1 as a prime – it isn't ordinarily considered to be prime.
$$1, 1, 2, 6, 30, 30, 210, \cdots$$

Recall that **prime numbers** are numbers greater than or equal to 2 which have only one and themselves as divisors. A **composite number** has factors other than itself and 1.

The primes (**A000040**) are
$$2, 3, 5, 7, 11, 13, 17, 19, \cdots$$

So for $n = 8$, the primordial number would be the product of all the primes less than 8, or

$$2 \cdot 3 \cdot 5 \cdot 7 = 210$$

Exercise: Find the first 15 terms of **A003418.**

I have saved the best for last. 1, 1, 2, 6 is also part of sequence **A001349** – the number of connected graphs with n nodes. We are back to graph theory!
A **connected graph** is a graph in which it is possible to move from any vertex to any other vertex along existing edges. Imagine the map of the United States with each state represented by a point, or node. Place an edge connecting two states if those states share any border – that is, if they are contiguous with each other. This

graph will not be a connected graph. Hawaii, for instance, will be an **isolated vertex** – a vertex with no edges joining it to any other vertex. It will not be possible to move from any other state to Hawaii along existing edges.

Let's draw all the connected graphs with three nodes.
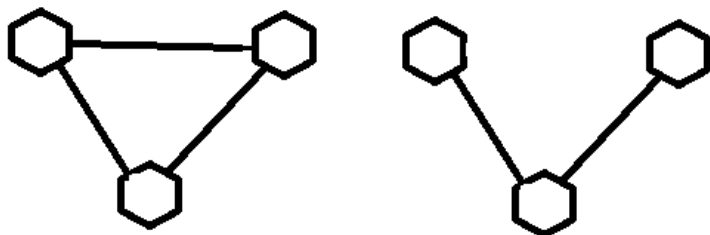
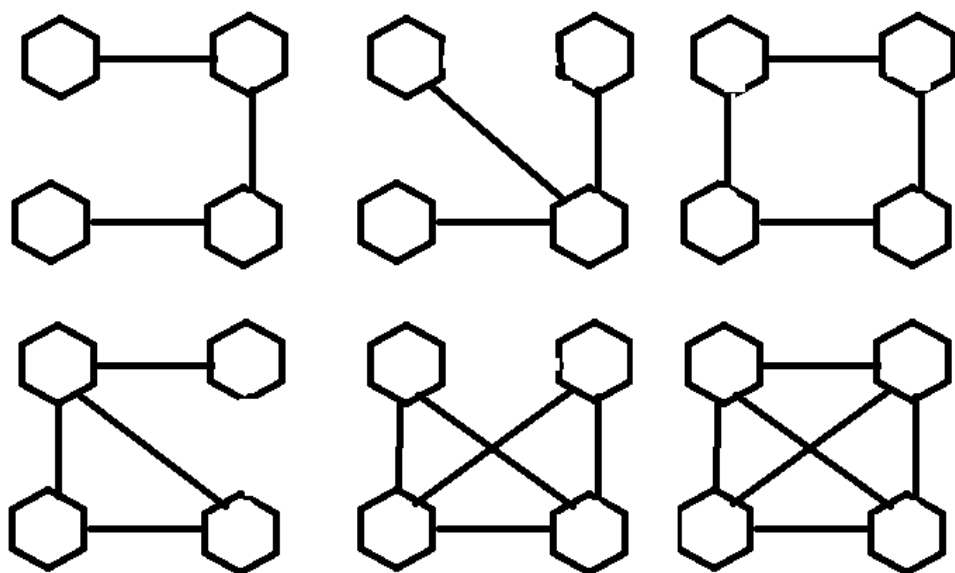
There are two of them.

Exercise: Draw all connected graphs with four nodes.

Answer:

Exercise: Draw all connected graphs with five nodes.

Even More Fun

Here are some challenges. Hello, Bill Mullan! Try to figure them out before looking them up.

A029579   "An obvious mixture of two sequences"

$$1, 1, 2, 3, 3, 5, 4, 7, 5, \cdots$$

A002808 $4, 6, 8, 9, 10, 12, 14, 15, \cdots$

A123865 $0, 15, 80, 255, \cdots$

Conclusion

We have seen how exploring the Online Encyclopedia of Integer Sequences can be a lot of fun, and also how it can lead us to many mathematical explorations. Seeing connections between things, solving problems, looking at familiar things in new ways – all of these help to develop our mathematical intuition. Intuition is the mental faculty of seeming to understand things or know answers or predict outcomes without apparently having all the needed facts. What's really going on is that your brain is using all kinds of knowledge and experience that you have begun to take for granted, so you aren't entirely aware you are using it.

Just starting with random sequences, we have explored number theory and graph theory. We've drawn pictures and made charts. We have proved theorems and discovered some mathematicians and the work they do. I have included a glossary of mathematical terms defined in this article at the end.

I hope you will continue this exploration. There is a fun Wikipedia game: give every player the same starting point, like an article on South American beetles, and see how many clicks it takes them to find a Wikipedia article on some completely different target subject, like the novel Frankenstein. I would like to suggest an OEIS version of the game. Start with any random string of integers – if it isn't in the OEIS you can pick a second one – and by clicking on linked sequences, try to eventually reach the prime numbers, the Fibonacci numbers, or some other sequence.

I can't resist adding one more. I just typed in

$$1, 2, 5, 9, 17$$

and I learn that this is the number of partitions of n if there are two kinds of 1 and two kinds of two! Now I will grab a blue marker and a red marker…

Glossary

Adjacency Matrix
Closed walk
Connected graph
Edge
Ellipsis
Factorial
Fibonacci number
Graph
Harmonic sequence
Isolated vertex
Node
Padovan number
Pancake Sequence
Partition
Prime number
Primorial Number

More Bonus Questions

1. Draw a digraph based on this adjacency matrix.

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

   Now find the sequence of closed walks starting at A of length 1, 2, 3, 4, 5, 6, etc.
   (There will be many zeros in this sequence.) Is it in the OEIS?

2. Can every positive integer be expressed as the sum of one or more Fibonacci numbers?

3. Suppose we define an odd partition as a partition including only odd integers. Find the first ten terms of the sequence of odd partitions of 1,2,3,…,10. Is this in the OEIS? What else is it equal to?

4. How many squares do you see?
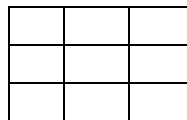
   This is called a 3×3 chessboard. Find the number of squares in an $n \times n$ chessboard for $n = 1, 2, 3, 4, 5$. Check your sequence in the OEIS.

5. Starting in the bottom left corner of a 3×3 chessboard and making only three kinds of moves – right one space, or diagonally up and over one space, or diagonally down and over – count how many ways there are to reach the bottom right corner. Do the same for $n \times n$ chessboards for $n = 1, 2, 3, 4, 5$. This is called the Motzkin sequence

Allowable Moves For The Motzkin Sequence.

6.   Read my other article and explore arithmetic derivatives in the OEIS.

References.

N. J. A. Sloane, editor, *The On-Line Encyclopedia of Integer Sequences*, published electronically at https://oeis.org,

Zeros and Multiplicity

Jeff VanHamlin

The path of the curve for the function $f(x)$ about its zeros depends on the multiplicity of each zero.

Theorem: If $f(x) = (x - c)^n \cdot g(x)$ where $g(x)$ is a polynomial of degree $n \geq 1$ and $g(c) \neq 0$.

Then $x = c$ is a zero of multiplicity n and the following holds:

    I.   If $n = 1$, the curve $f(x)$ crosses the x-axis at $x = c$ in a non-horizontal linear path.
    II.   If $n = 2$, the curve $f(x)$ has a local minimum or maximum on the x-axis at $x = c$.
    III.   If $n = 3$, the curve $f(x)$ has a saddle point on the x-axis at $x = c$.

In general; for $n > 1$ and even, the function has a local minimum or maximum at $x = c$ and for $n > 1$ and odd, the function has a saddle point at $x = c$.

Proof:

Let $f(x) = (x - c)^n \cdot g(x)$, where $g(x)$ is any polynomial where $g(c) \neq 0$ and $n \in \mathbb{N}$. Therefore $f(x)$ is a polynomial which has a zero of $x = c$ with multiplicity n.

$$f'(x) = n(x - c)^{n-1} \cdot g(x) + (x - c)^n \cdot g'(x)$$

$$= (x - c)^{n-1}[n \cdot g(x) + (x - c) \cdot g'(x)]$$

    I.   If $n = 1$, then

$$f'(x) = (x - c)^{1-1}[1 \cdot g(x) + (x - c) \cdot g'(x)]$$

$$= g(x) + (x - c) \cdot g'(x)$$

Therefore at $x = c$

$$f'(c) = g(c) = non - zero\ constant$$

This means around the zero, the slope is constant, and therefore linear. Moreover, the slope of the linear path around the zero is $g(c)$, since the tangent line to the curve at $x = c$ is $y = g(c)(x - c)$.

    II.   If $n = 2$, then

$$f'(x) = (x - c)^{2-1}[2 \cdot g(x) + (x - c) \cdot g'(x)]$$

$$= (x - c)[2 \cdot g(x) + (x - c) \cdot g'(x)]$$

Therefore at $x = c$

$$f'(c) = 0$$

This means at $x = c$ the, the slope of the curve is 0. Giving a tangent line of $y = 0$ at $x = c$.

By taking the second derivative, we can check the concavity of the curve.

$$f''(x) = 1 \cdot [2g(x) + (x - c)g'(x)] + (x - c)[2g'(x) + g'(x) + (x - c)g''(x)]$$

At $x = c$,

$$f''(c) = 2g(c) = non - zero\ constant$$

Therefore, in the neighborhood around the $x = c$, $f(x)$ is either concave up or concave down depending on whether $g(c)$ is positive or negative. Therefore, the function $f(x)$ has a local minimum or local maximum at $x = c$.

III.   If $n = 3$, then

$$f'(x) = (x - c)^{3-1}[3 \cdot g(x) + (x - c) \cdot g'(x)]$$

$$= (x - c)^2[3 \cdot g(x) + (x - c) \cdot g'(x)]$$

Therefore at $x = c$

$$f'(c) = 0$$

This means at $x = c$ the, the slope of the curve is 0. Giving a tangent line of $y = 0$ at $x = c$.

By taking the second derivative, we can check concavity of the curve.

$$f''(x) = 2(x - c) \cdot [3g(x) + (x - c)g'(x)] + (x - c)^2[3g'(x) + g'(x) + (x - c)g''(x)]$$

At $x = c$,

$$f''(c) = 0$$

To verify that $x = c$ is a saddle point and not a local minimum, it must be verified that $f'(x)$ has the same sign before and after $x = c$.

Since

$$f'(x) = (x - c)^2[3 \cdot g(x) + (x - c) \cdot g'(x)]$$

$$= (x - c)^2[h(x)]$$

Then $x = c$ is a local minimum of $f'(x)$, verifying that the function is monotonic and therefore increasing or decreasing in the neighborhood of $x = c$.

Therefore $x = c$ is a saddle point of $f(x)$.

Next, we will prove the factors of even multiplicity are local minimums or maximums at the zero and factors of odd multiplicity greater than 1 have a saddle point at the zero.

An induction proof will be used. We will first prove the even multiplicities and then the odd.

In general

$$f(x) = (x - c)^n \cdot g(x)$$

$$f'(x) = n(x - c)^{n-1} \cdot g(x) + (x - c)^n \cdot g'(x)$$

$$= (x - c)^{n-1}[n \cdot g(x) + (x - c)g'(x)]$$

$$f''(x) = n(n-1)(x - c)^{n-2} \cdot g(x) + n(x - c)^{n-1}g'(x) + n(x - c)^{n-1} \cdot g'(x) + (x - c)^n \cdot g''(x)$$

$$= (x - c)^{n-2}[n(n-1) \cdot g(x) + 2n(x - c) \cdot g'(x) + (x - c)^2 \cdot g''(x)]$$

What now needs to be proved is the shape of the curve around the zero for even and odd multiplicities for integer values greater than $n \geq 1$.

Let n be an even number greater than 1.

For $n = 2$, it was verified that the curve has a local minimum or maximum at $x = c$.

As for the induction proof we will assume the function

$$y = (x - c)^n \cdot g(x)$$

has a local minimum or maximum at $x = c$ for any even interger n

Therefore, we must show

$$f(x) = (x - c)^{n+2} \cdot g(x)$$

has a local minimum or maximum at $x = c$, since $(x - c)^{n+2}$ is the next zero of even multiplicity.

The function has a zero at $x = c$, since

$$f(c) = 0$$

The function has a horizontal tangent line, because

$$f'(x) = (n + 2)(x - c)^{n+1} \cdot g(x) + (x - c)^{n+2} \cdot g'(x)$$

$$= (x - c)^{n+1}[(n + 2) \cdot g(x) + (x - c) \cdot g'(x)]$$

And at $x = c$

$$f'(c) = 0$$

Finally

$$f''(x) = (n + 2)(n + 1)(x - c)^n \cdot g(x) + (n + 2)(x - c)^{n+1} \cdot g'(x) + (n + 2)(x - c)^{n+1} \cdot g'(x) + (x - c)^{n+2} \cdot g''(x)$$

$$= (x - c)^n[(n + 2)(n + 1) \cdot g(x) + 2(x + 2)(x - c) \cdot g'(x) + (x - c)^2 \cdot g''(x)]$$

At $x = c$,

$$f''(c) = 0$$

Notice that the second derivative has the form

$$f''(x) = (x - c)^n \cdot [h(x)]$$

Therefore, from the assumption, the factor $x - c$ has even multiplicity. So $f''(x)$ has a local minimum or maximum at $x = c$ and thus is concave up or concave down at $x = c$, so the first derivative changes sign before and after the zero $x = c$. Therefore $f(x)$ has a local minimum or local maximum at $x = c$.

Now to prove the odd multiplicity for n greater than 1.

Let n be an odd number greater than 1.

It has been shown that when $n = 3$, the function has a saddle point at $x = c$.

Assume

$$y = (x - c)^n \cdot g(x)$$

To complete the induction proof, it must be shown that the function

$$f(x) = (x - c)^{n+2} \cdot g(x)$$

has a saddle point at $x = c$, since $(x - c)^{n+2}$ is the next zero of odd multiplicity.

The function has a zero at $x = c$, since

$$f(c) = 0$$

The function has a horizontal tangent line, because

$$f'(x) = (n + 2)(x - c)^{n+1} \cdot g(x) + (x - c)^{n+2} \cdot g'(x)$$

$$= (x - c)^{n+1}[(n + 2) \cdot g(x) + (x - c) \cdot g'(x)]$$

And at $x = c$

$$f'(c) = 0$$

Finally

$$f''(x) = (n + 2)(n + 1)(x - c)^n \cdot g(x) + (n + 2)(x - c)^{n+1} \cdot g'(x) + (n + 2)(x - c)^{n+1} \cdot g'(x) + (x - c)^{n+2} \cdot g''(x)$$

$$= (x - c)^n[(n + 2)(n + 1) \cdot g(x) + 2(x + 2)(x - c) \cdot g'(x) + (x - c)^2 \cdot g''(x)]$$

At $x = c$,

$$f''(c) = 0$$

The second derivative has the form

$$f''(x) = (x - c)^n \cdot [h(x)]$$

Since n is an odd number, $f''(x)$ has a saddle point at $x = c$. Therefore the concavity of $f''(x)$ changes at the zero $x = c$, so $f'(x)$ has the same sign before and after $x = c$. So the function $f(x)$ is monotonic around the zero $x = c$, and therefore a saddle point.

Finding The Tangent Line to a Rational Function, using Division

Nick Goins

In this article, I will present a method for finding a tangent line to the graph of a rational function, using polynomial division. As we will see, the method will apply only to a certain sub-collection of rational functions and it will require that the point of tangency be on the y-axis. We will then discuss conditions under which we can generalize the method. That is, we will consider points of tangency which are not on the y-axis, and we will see when we are able to use a horizontal shift so that the corresponding point is on the y-axis. To motivate the method, we will begin with an example.

Example 1: Consider the rational function $f(x) = \frac{x^3+x^2-x+4}{x^2+1}$. Using polynomial division to write $f(x)$ as a proper rational function, we get

$$f(x) = x + 1 + \frac{-2x + 3}{x^2 + 1}$$

Now find the tangent line to the graph of $f(x)$ at $x = 0$. The point of tangency is $(0, f(0)) = (0, 4)$ and the derivative of $f(x)$ is

$$f'(x) = \frac{(3x^2 + 2x - 1)(x^2 + 1) - (x^3 + x^2 - x + 4)2x}{(x^2 + 1)^2}$$

so that the slope of the tangent line is $m = f'(0) = \frac{-1 \cdot 1 - 0}{1} = -1$. The tangent line is

$$y - y_1 = m(x - x_1) \qquad \Rightarrow \qquad y - 4 = -1(x - 0) \qquad \Rightarrow \qquad y = -x + 4$$

Notice that if we add the quotient and remainder found from using polynomial division, we get

$$(x + 1) + (-2x + 3) = -x + 4$$

□

That is, the tangent line is the same as the sum of the quotient and the remainder. Is this a very unique case or is there an underlying reason for this to happen more generally?

Note: Before answering the question in general, let's consider the following function similar to the above example,

$$h(x) = mx + b + \frac{cx + d}{\alpha x^2 + \beta x + \gamma}$$

We want to see if this function has the property that the tangent line to the graph of $h(x)$ at $x = 0$ is the sum of $mx + b$ and $cx + d$. We also want to see if there are any requirements on the coefficients for this to happen. The point of tangency for $h(x)$ at $x = 0$ is

$$(0, h(0)) = \left(0, b + \frac{d}{\gamma}\right)$$

The derivative of $h(x)$ is

$$h'(x) = m + \frac{c(\alpha x^2 + \beta x + \gamma) - (cx + d)(2\alpha x + \beta)}{(\alpha x^2 + \beta x + \gamma)^2}$$

and the slope of the tangent line is

$$h'(0) = m + \frac{c\gamma - d\beta}{\gamma^2}$$

The equation of the tangent line is

$$y - y_1 = m(x - x_1) \qquad \Rightarrow \qquad y - \left(b + \frac{d}{\gamma}\right) = \left(m + \frac{c\gamma - d\beta}{\gamma^2}\right)(x - 0)$$

which reduces to

$$y = mx + b + \left(\frac{c\gamma - d\beta}{\gamma^2}\right)x + \frac{d}{\gamma}$$

The first two terms are the quotient and now we will set the last two terms equal to the remainder to find conditions on the coefficients. That is,

$$\left(\frac{c\gamma - d\beta}{\gamma^2}\right)x + \frac{d}{\gamma} = cx + d \qquad \Rightarrow \qquad \gamma = 1, \beta = 0$$

That is, the result holds if the denominator has the form $\alpha x^2 + 1$. The following theorem provides the general answer to the question posed above.

Theorem 1: For a rational function $f(x) = \frac{g(x)}{d(x)}$, the tangent line to $f(x)$ at $x = 0$ will be identical to the sum of the quotient and remainder if and only if $f(x)$ can be written in the following form

$$f(x) = m_1 x + b_1 + \frac{r(x)}{d(x)}$$

where the remainder, $r(x)$, is linear and $d(x)$ is a polynomial of degree at least two, and has the form $d(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_2 x^2 + 1$.

Proof: The point of tangency is

$$(0, f(0)) = \left(0, b + \frac{r(0)}{d(0)}\right)$$

To find the slope of the tangent line, we need to differentiate $f(x)$

$$f'(x) = m + \frac{r'(x)d(x) - r(x)d'(x)}{(d(x))^2}$$

so that the slope of the tangent line is

$$m = f'(0) = m + \frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2}$$

Then, the equation of the tangent line is

$$y - \left(b + \frac{r(0)}{d(0)}\right) = \left(m + \frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2}\right)(x - 0)$$

$$y = b + \frac{r(0)}{d(0)} + mx + \left(\frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2}\right)x$$

$$y = mx + b + \left(\frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2}\right)x + \frac{r(0)}{d(0)}$$

We want to find conditions on $r(x)$ and $d(x)$ so that $\left(\frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2}\right)x + \frac{r(0)}{d(0)}$ is equal to the remainder. Since the function $r(x)$ is linear, we can write $r(x) = m_2 x + b_2$. For this to happen we need to solve the following system of equations

$$\begin{cases} \dfrac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2} = m_2 \\[2em] \dfrac{r(0)}{d(0)} = b_2 \end{cases}$$

From the definition of $r(x)$, we get $r(0) = b_2$, so that the second equation above gives $d(0) = 1$. Substituting this into the first equation in the system gives

$$\frac{r'(0)d(0) - r(0)d'(0)}{(d(0))^2} = m_2 \qquad \Rightarrow \qquad m_2 - b_2 d'(0) = m_2$$

This last equation then says that $d'(0) = 0$. Knowing that $d(x)$ is a polynomial with $d(0) = 1$, we know that the constant term must be 1. Knowing that $d'(0) = 0$, we know that the coefficient of the first degree term in $d(x)$ must be 0. That is, the form of $d(x)$ must be $a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + 1$.

□

Example 2: The tangent line to $f(x) = 3x + 5 + \frac{2x-1}{x^3+4x^2+1}$, at $x = 0$ is $y = 5x + 4$.

□

Corollary: The following is a direct result of the theorem.

$$f(x) = m_1 x + b_1 + \frac{m_2 x + b_2}{a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + 1} \qquad \Rightarrow \qquad f'(0) = m_1 + m_2$$

Example 4: Show that $x = 0$ is a critical point of $f(x) = \frac{x^3+1}{x^2+2}$, using only algebra.

Solution: So that the function has the form in the corollary, we can factor out the 2 from the denominator,

$$f(x) = \frac{x^3 + 1}{x^2 + 2} = \frac{1}{2}\left(\frac{x^3 + 1}{\frac{1}{2}x^2 + 1}\right)$$

Then, consider the function $g(x) = \frac{x^3+1}{\frac{1}{2}x^2+1}$, so that $f(x) = \frac{1}{2}g(x)$ and therefore $f'(x) = \frac{1}{2}g'(x)$. Using polynomial division on $g(x)$ gives

$$g(x) = 2x + \frac{-2x + 1}{\frac{1}{2}x^2 + 1}$$

From the corollary, we get $g'(0) = 2 - 2 = 0$ and therefore, $f'(0) = \frac{1}{2} \cdot 0 = 0$.

$\square$

A Slight Generalization of the Above Technique

Consider the function $f(x) = (x - 3)^2$. The point of tangency for the tangent line at $x = 2$ is $(2, f(2)) = (2,1)$ and the slope is $m = f'(2) = 2(2 - 3) = -2$. Thus, the tangent line is $y - 1 = -2(x - 2)$, which is $y_1(x) = -2x + 5$. Now, consider shifting the graph of $f(x)$ so that the point of tangency moves to the y-axis. To accomplish this, the point $(2,1)$ shifts two units to the left to the point $(0, 1)$. That is, we consider the function $g(x) = f(x + 2)$. Now find the tangent line to $g(x)$ at $x = 0$. The function expression for $g(x)$ is $g(x) = (x - 1)^2$ and the point of tangency is $(0, g(0)) = (0, 1)$ and the slope of the tangent line is $m = g'(0) = 2(0 - 1) = -2$ so that the tangent line is $y - 1 = -2(x - 0)$, giving $y_2(x) = -2x + 1$. Finally, consider shifting the tangent line $y_2(x)$ two units to the right,

$$y_2(x - 2) = -2(x - 2) + 1 = -2x + 5$$

The idea in the above discussion will be stated and as a theorem and proved next. We will then return to the idea of finding tangent lines to rational functions, in particular at locations off of the y-axis.

Theorem 2 (Horizontal Shift of Tangent Lines): If $g(x) = f(x + c)$ and $y_1(x)$ is the tangent line to $g(x)$ at $x = 0$ then $y_2(x) = y_1(x - c)$ is the tangent line to $f(x)$ at $x = c$.

Proof: Suppose $y_1(x)$ is the tangent line to $g(x)$ at $x = 0$. Then,

$$y_1 = m_1(x - x_1) + b_1 \qquad \Rightarrow \qquad y_1 = g'(0)(x - 0) + g(0)$$

Since $g(x) = f(x + c)$, then $g(0) = f(c)$ and $g'(x) = f'(x + c)$, so that $g'(0) = f'(0 + c) + f'(c)$,

$$\Rightarrow \qquad y_1 = f'(c)x + f(c)$$

Then, $y_2(x) = y_1(x - c) = f'(c)(x - c) + f(c)$, which is equivalent to

$$y_2 - f(c) = f'(c)(x - c)$$

This latter equation is the tangent line to $f(x)$ at $x = c$.

□

Example 5: Find the tangent line to $f(x) = \frac{2x^3 + x - 1}{(x-2)^2 + 1}$ at $x = 2$, without computing a derivative.

Solution: Since the point of tangency is not on the y-axis, we will shift the graph horizontally to the left so that the point of tangency will be on the y-axis. We will find the tangent line for the shifted function, then perform the opposite shift to the tangent line to obtain the tangent line to the original graph. That is, define $g(x) = f(x + 2)$. Then, the expression for $g(x)$ is

$$g(x) = f(x + 2) = \frac{2(x + 2)^3 + (x + 2) - 1}{((x + 2) - 2)^2 + 1} = \frac{2x^3 + 12x^2 + 25x + 17}{x^2 + 1}$$

Using long division to write the expression as a proper rational function gives,

$$g(x) = 2x + 12 + \frac{23x + 5}{x^2 + 1}$$

Then, the tangent line to $g(x)$ at $x = 0$ is

$$y_2(x) = (2x + 12) + (23x + 5) = 25x + 17$$

To obtain the tangent line to the original function $f(x)$, at $x = 2$, we will shift the tangent line y_2 two units to the right. That is,

$$y_1(x) = y_2(x - 2) = 25(x - 2) + 17 = 25x - 33$$

□

As the previous example illustrates, we can apply the method of finding a tangent line to a rational function using the results of division, if the rational function has the form $f(x) = \frac{p(x)}{(x-c)^2 + 1}$, where the degree of $p(x)$ is 3. The most general case is the following

$$f(x) = \frac{p(x)}{d(x - c)}$$

where, the degree of $p(x)$ is one more than the degree of $d(x)$ and $d(x)$ has the form $a_k x^k + a_{k-1} x^{k-1} + \cdots + a_2 x^2 + 1$.

In the following theorem, note that the first use of $x = h$ refers to a specific x-value whereas the second use refers to a vertical line.

Theorem 3: For a rational function $f(x) = \frac{p(x)}{d(x)}$, such that $\deg(p(x)) = 3$ and $d(x)$ has no real roots, we can

find the tangent line to $f(x)$ at $x = h$, where $x = h$ is the axis of symmetry of $d(x)$, using only algebra.

Proof: Writing $d(x)$ in standard form, and using the fact that it has no real roots, we can write the following

$$f(x) = \frac{p(x)}{d(x)} = \frac{p(x)}{a(x-h)^2 + k} = \frac{1}{k}\left[\frac{p(x)}{\frac{a}{k}(x-h)^2 + 1}\right]$$

$\square$

Example 6: For the function, $f(x) = \frac{x^3+1}{x^2+4x+5}$, we can write

$$f(x) = \frac{x^3 + 1}{x^2 + 4x + 5} = \frac{x^3 + 1}{(x+2)^2 + 1}$$

so that the tangent line can be found at $x = -2$, by first doing a horizontal shift. Similarly, consider the rational function $g(x) = \frac{x^3+1}{x^2+4x+7}$, where we can write

$$g(x) = \frac{x^3 + 1}{x^2 + 4x + 7} = \frac{x^3 + 1}{(x+2)^2 + 3} = \frac{1}{3}\left(\frac{x^3 + 1}{\frac{1}{3}(x+2)^2 + 1}\right)$$

For this function we can also find the tangent line at $x = -2$.

$\square$

If we have a tangent line to a function $g(x)$ at $x = a$ and we know that $g(x)$ and $f(x)$ are related by $g(x) = cf(x)$ then how do we obtain the tangent line to $f(x)$? The following theorem will state and prove the fact that if we stretch or compress a function, then we correspondingly stretch, or compress, the tangent lines by the same factor.

Theorem 4: If $g(x) = cf(x)$ and if y_t is the tangent line to $f(x)$ at $x = a$, then cy_t is the tangent line to $g(x)$ at $x = a$.

Proof: The tangent line to $f(a)$ at $x = a$, is $y_t = f'(a)(x - a) + f(a)$. The point of tangency on the graph of $g(x)$ is
$(a, g(a))$ and the slope is $m = g'(a)$, but with the relationship $g(x) = cf(x)$, we can write this information as $(a, cf(a))$ and $m = cf'(a)$, respectively. Thus, the tangent line to $g(x)$ at $x = a$ is

$$y - cf(a) = cf'(a)(x-a) \quad \Rightarrow \quad y = cf'(a)(x-a) + cf(a) \quad \Rightarrow \quad y$$
$$= c(f'(a)(x-a) + f(a))$$

$\square$

The following example will illustrate Theorem 3 and Theorem 4.

Example 7: Let $f(x) = \frac{x^3+x^2+x+1}{x^2+6x+13}$. Find the tangent line to $f(x)$ at $x = h$, where $x = h$ is the axis of symmetry of the denominator.

Solution: Completing the square on the denominator gives the following for $f(x)$

$$f(x) = \frac{x^3 + x^2 + x + 1}{(x+3)^2 + 4}$$

From the previous theorem, we see that we can find the tangent line to this function at $x = -3$, by first shifting the graph three units to the right. That is, we will define $g(x) = f(x - 3)$ and find the tangent line to $g(x)$ at $x = 0$.

$$g(x) = f(x - 3) = \frac{(x - 3)^3 + (x - 3)^2 + (x - 3) + 1}{((x - 3) + 3)^2 + 4}$$

$$= \frac{x^3 - 8x^2 + 22x - 20}{x^2 + 4}$$

Factoring out the 4 from the denominator puts the function expression in the form we need,

$$= \frac{1}{4} \left[\frac{x^3 - 8x^2 + 22x - 20}{\frac{1}{4}x^2 + 1} \right]$$

Define $h(x) = \frac{x^3 - 8x^2 + 22x - 20}{\frac{1}{4}x^2 + 1}$, so that $g(x) = \frac{1}{4}h(x)$. Then, using polynomial division on $h(x)$ gives

$$h(x) = 4x - 32 + \frac{18x + 12}{\frac{1}{4}x^2 + 1}$$

Thus, the tangent line to $h(x)$ at $x = 0$ is $y = (4x - 32) + (18x + 12) = 22x - 20$. By a previous theorem, the tangent line to $g(x)$ at $x = 0$ is $y = \frac{1}{4}(22x - 20)$ and finally, by shifting three units to the left, the tangent line to $f(x)$ at $x = 3$ is $y = \frac{1}{4}(22(x + 3) - 20) = \frac{1}{4}(22x + 66 - 20) = \frac{11}{2}x + \frac{23}{2}$.

□

In all of the preceding discussion, we focused on rational functions and certain values of the first derivative, the following theorem provides a basic result regarding a class of rational functions and a value of the second derivative.

Theorem 5: For a rational function of the form,

$$f(x) = m_1 x + b_1 + \frac{m_2 x + b_2}{a_k x^k + a_{k-1}x^{k-1} + \cdots + a_3 x^3 + 1}$$

$x = 0$ will be a zero of f'', but is not an inflection point.

Proof: To simplify the notation, we will write $f(x) = m_1 x + b_1 + \frac{r(x)}{d(x)}$, where $r(x)$ and $d(x)$ are the functions stated above. The first derivative of $f(x)$ is

$$f'(x) = m_1 + \frac{r'd - rd'}{d^2}$$

Since $r(x)$ is a linear function, we know that its derivative is m_2. Substituting this is and taking the second derivative of $f(x)$ gives

$$f'(x) = m_1 + \frac{m_2 d - rd'}{d^2} \qquad \Rightarrow \qquad f''(x) = \frac{(m_2 d' - m_2 d' - rd'')d^2 - (m_2 d - rd')2dd'}{d^4}$$

Then, the second derivative evaluated at $x = 0$ is

$$f''(0) = \frac{(m_2 d'(0) - m_2 d'(0) - r(0)d''(0))d^2(0) - (m_2 d(0) - r(0)d'(0))2d(0)d'(0)}{d^4(0)}$$

We know the function values, $r(0) = b_2, d(0) = 1, d'(0) = 0$ and $d''(0) = 0$, so that the above reduces to

$$f''(0) = \frac{-b_2 d''(0)}{d^4(0)} = 0$$

This shows that $x = 0$ is a potential inflection point, but it will depend on the sign of $f''(x)$ on either side of $x = 0$. Since the numerator is a polynomial, we can determine its degree. Also, since the denominator will be strictly positive, the sign is completely determined by the numerator. Consider the following function

$$y = -rd''d^2 - 2dd'(m_2 d - rd')$$

Let $\deg(d(x))$ denote the degree of the function $d(x)$. Since $r(x)$ is linear, $\deg(r(x)) = 1$. Then, the degree of the first term is

$$(\deg(d(x)) - 2) \cdot 2\deg(d(x))$$

which reduces to

$$2(\deg(d(x)))^2 - 4\deg(d(x))$$

That is, the first term has even degree. The degree of the second term is

$$\deg(d(x)) \cdot (\deg(d(x)) - 1) \cdot \deg(d(x))$$

which reduces to

$$(\deg(d(x)))^3 - (\deg(d(x)))^2$$

which is also even. Thus, the numerator is an even degree polynomial and so does not change sign on either side of its zero.

Composition of piecewise functions and their domain

Jeff VanHamlin

I am going to present a method for determining the function expression for the composition of piecewise functions. To do this, lets first look at the case of evaluating a composition of piecewise functions at a value. Consider the piecewise functions

$$f(x) = \begin{cases} 4x - 3, & x < -1 \\ -2x + 1, & x \geq -1 \end{cases}$$

and

$$g(x) = \begin{cases} -x - 7, & x < 3 \\ 5x + 2, & x \geq 3 \end{cases}$$

When evaluating a piecewise function at an input, the input restrictions determine which of the functions we are to use in the piecewise expression. If we want to evaluate the composition $(f \circ g)(4)$, we must first evaluate $g(4)$ and then evaluate the function f at the output value of $g(4)$. To do this we first determine which set or restriction the element 4 belongs to in the g function. Since 4 satisfies the inequality $x \geq 3$, it belongs to that restriction set. Therefore, we use the corresponding function rule $5x + 2$, giving $5(4) + 2 = 22$. Now we must evaluate $f(22)$. To do this we determine which restriction set that 22 belongs to in the f function. Since 22 satisfies the inequality $x \geq -1$, we will use the corresponding function rule $-2x + 1$. This will give us a final answer of $-2(22) + 1 = -43$.

This can be condensed into the calculation below

$$(f \circ g)(4) = f\big(g(4)\big) = f(5(4) + 2) = f(22) = -2(22) + 1 = -43$$

This gives us some insight into building a function rule for the composition. The composition depends largely on the restrictions. To evaluate an input, x, for the composition we first look at the restrictions on the g function. We will use the corresponding function rule depending on what restriction set the value of x satisfies. To determine what function rule of f is used after the evaluation of g, we look at the set the output value $g(x)$ satisfies.

The above calculation of $f\big(g(4)\big)$ used the function rule was $-2(5x + 2) + 1$ evaluated at $x = 4$. We used the composition of the $g(x) = 5x + 1$ with $f(x) = -2x + 1$ because $x = 4$ is in the restriction set of $x \geq 3$ and $g(4) = 22$ is in the restriction set of $x \geq -1$. So we must look at the intersection of the sets.

Function expression for the composition of piecewise functions

Using the same functions, we determine the composition expression of the piecewise functions. Given

$$f(x) = \begin{cases} 4x - 3, & x < -1 \\ -2x + 1, & x \geq -1 \end{cases}$$

and

$$g(x) = \begin{cases} -x - 7, & x < 3 \\ 5x + 2, & x \geq 3 \end{cases}$$

To determine the composition, we will start with the restriction sets of $g(x)$.

Starting with the top restriction set of the g function. If $x < 3$, there are two possible restriction outcomes of the function rule $g(x) = -x - 7$. Either the outputs will be $g(x) < -1$ or $g(x) \geq -1$. This means we must solve these inequalities for both of piecewise expressions of g.

Solving $g(x) < -1$, gives $-x - 7 < -1 \implies x > -6$.

Since we are looking at the x-values that satisfy $x < 3$, we will use the expression $-x - 7$ in g. Also since $g(x) < -1$, we will use the expression $4x + 3$ in f. Therefore, we will compose these two expressions at the intersection of the sets $x < 3$ and $x > -6$. Resulting in the composition

$$4(-x - 7) + 3, \qquad -6 < x < 3$$

Solving $g(x) \geq -1$, gives $-x - 7 \geq -1 \implies x \leq -6$.

Since we are looking at the x-values that satisfy $x < 3$, we will use the expression $-x - 7$ in g. Also since $g(x) \geq -1$, we will use the expression $-2x + 1$ in f. Therefore, we will compose these two expressions at the intersection of the sets $x < 3$ and $x \leq -6$. Resulting in

$$-2(-x - 7) + 1, \qquad x \leq -6$$

Now we must repeat this process for the next restriction set of g. If $x \geq 3$, again there are two possible restriction outcomes for the expression $g(x) = 5x + 2$. Either the outputs will be $g(x) < -1$ or $g(x) \geq -1$. This means we must solve these inequalities for both of piecewise expressions of g.

Solving $g(x) < -1$, gives $5x + 2 < -1 \implies x < -\frac{3}{5}$.

Since we are looking at the x-values that satisfy $x \geq 3$, we will use the expression $5x + 2$ in g. Also since $g(x) < -1$, we will use the function expression $4x + 3$ in f. Therefore, we will compose these two expressions at the intersection of the sets $x \geq 3$ and $x < -\frac{3}{5}$. Notice that the intersection of these sets is the empty set. Therefore, the composed expression $4(5x + 2) + 3$ will not be used.

Solving $g(x) \geq -1$, gives $5x + 2 \geq -1 \implies x \geq -\frac{3}{5}$.

Since we are looking at the x-values that satisfy $x \geq 3$, we will use the expression $5x + 2$ in g. Also since $g(x) \geq -1$, we will use the expression $-2x + 1$ in f. Therefore, we will compose these two expressions at the intersection of the sets $x \geq 3$ and $x \geq -\frac{3}{5}$. Resulting in

$$-2(5x + 2) + 1, \qquad x \geq 3$$

The final expression for the composition becomes

$$(f \circ g)(x) = \begin{cases} -2(-x-7)+1, & x \leq -6 \\ 4(-x-7)+3, & 6 < x < 3 \\ -2(5x+2)+1, & x \geq 3 \end{cases}$$

$$= \begin{cases} 2x+15, & x \leq -6 \\ -4x-25, & 6 < x < 3 \\ -10x-3, & x \geq 3 \end{cases}$$

To generalize this technique. Given $i, j, n, m \in \mathbb{N}$, $0 \leq i \leq m$, $0 \leq j \leq n$ and A_i is the domain of f_i and B_i is the domain of g_j. Given the functions

$$f(x) = \begin{cases} f_1, & x \in A_1 \\ f_2, & x \in A_2 \\ \vdots \\ f_m, & x \in A_m \end{cases}$$

and

$$g(x) = \begin{cases} g_1, & x \in B_1 \\ g_2, & x \in B_2 \\ \vdots \\ g_n, & x \in B_n \end{cases}$$

Then the composition of the functions is

$$(f \circ g)(x) = \begin{cases} f_1 \circ g_1, & x \in B_1 \cap g(x) \in A_1 \\ f_1 \circ g_2, & x \in B_2 \cap g(x) \in A_1 \\ \vdots \\ f_1 \circ g_n, & x \in B_n \cap g(x) \in A_1 \\ f_2 \circ g_1, & x \in B_1 \cap g(x) \in A_2 \\ f_2 \circ g_2, & x \in B_2 \cap g(x) \in A_2 \\ \vdots \\ f_2 \circ g_n, & x \in B_n \cap g(x) \in A_2 \\ \vdots \\ f_m \circ g_1, & x \in B_1 \cap g(x) \in A_n \\ f_m \circ g_2, & x \in B_2 \cap g(x) \in A_n \\ \vdots \\ f_m \circ g_n, & x \in B_m \cap g(x) \in A_n \end{cases} = \{f_i \circ g_j, \ x \in B_j \cap g(x) \in A_i$$

If $x \in B_j \cap g(x) \in A_i$ is the empty set, exclude the composed function expression $f_i \circ g_j$ from the composition. This also means the amount of piecewise expressions in the composition will be at most the product of m and n.

Domain of a composition

If

$$f(x) = \begin{cases} \dfrac{x-2}{x+3}, & x < 0 \\ 3x + 1, & x \geq 0 \end{cases}$$

and

$$g(x) = \begin{cases} \sqrt{x+5}, & x < 1 \\ x^2 - 4, & x \geq 1 \end{cases}$$

The domain of a composition is defined by

$$Dom(f \circ g) = \{x \in Dom(g) \cap (Rng(g) \in Dom(f))\}$$

First, we must find the domain of g. This can be found by determining the domain of the piecewise expressions on their corresponding restrictions.

For the expression $\sqrt{x+5}$ we know that $x + 5 \geq 0$ on the restriction of $x < 1$. Therefore, the function rule $\sqrt{x+5}$ is used for the x-values at the intersection of $x \geq -5$ and $x < 1$ which is the set $[-5,1)$. Since the expression $x^2 - 4$ is defined for all x, we will use the expression $x^2 - 4$ for all values on the set $x > 1$ which is the set $[1, \infty)$. This means the domain of g is the set $[-5,1) \cup [1, \infty)$. Therefore $Dom(g) = [-5, \infty)$.

The domain of f is found similarly.

For the expression $\frac{x-2}{x+3}$, we know that $x + 3 \neq 0$, so $x \neq -3$ on the restriction $x < 1$. Since $x = -3$ is an element of the set $x < 1$ it must be excluded from the domain of f. This is the set $(-\infty, -3) \cup (-3,1)$ The second expression of f is $3x + 1$ and it is defined for all real numbers on the restriction $x \geq 1$. This is the set $[1, \infty)$. Therefore $dom(f) = (-\infty, -3) \cup (-3,1) \cup [1, \infty) = (-\infty, -3) \cup (-3, \infty)$.

Now we must determine when $g(x)$ is an element of the domain of f. This is done by determining when the output of the g function is an element of the domain of the f function.

$$Rng(g) \in Dom(f)$$

When determining $Rng(g) \in Dom(f)$ we investigate both g expressions $\sqrt{x+5}$ and $x^2 - 4$, we must look at the invalid input of -3 for the domain of f. We determine any value of x when evaluated g gives an output of -3 and exclude those values.

$$\sqrt{x+5} \neq -3 \quad \rightarrow \quad infinite\ solutions = (-\infty, \infty)$$

The expression $\sqrt{x+5}$ is always non-negative, giving the solution of all real numbers.

$x^2 - 4 \neq -3 \quad \rightarrow \quad x^2 \neq 1 \quad \rightarrow \quad x \neq \pm 1 = (-\infty, -1) \cup (-1, 1) \cup (1, \infty)$

When using the expression $\sqrt{x+5}$, we know it is restricted to the values $x \geq 1$, therefore we need only exclude $x = 1$ from the domain of the composition.

Note, $Dom(g) = [-5, \infty)$ and $Rng(g) \in Dom(f)$ when $x \in (-\infty, 1) \cup (1, \infty)$. The intersection of these two sets give the domain of the compostion.

Therefore $dom(f \circ g) = [-5, 1) \cup (1, \infty)$.

To verity this is correct, we will find the function expression for this composition and verify the domain found is correct.

Solve all $g(x)$ expressions for the restrictions of f and determine their intersections with the restrictions of g.

$$\sqrt{x+5} < 0 \quad \rightarrow \quad no\ solution$$

This means the composed expression $\frac{\sqrt{x+5}-2}{\sqrt{x+5}+3}$ will not be used.

$$\sqrt{x+5} \geq 0 \quad \rightarrow \quad x + 5 \geq 0 \quad \rightarrow \quad x \geq -5$$

Now the intersection of $x \geq -5$ and $x < 1$ is $-5 \leq x < 1$. Giving the composed expression $\quad 3\sqrt{x+5}+1$ on the restriction of $-5 \leq x < 1$.

$$x^2 - 4 < 0 \quad \rightarrow \quad -2 < x < 2$$

The intersection of $-2 < x < 2$ and $x \geq 1$ is $1 \leq x < 2$. Giving the composed expression

$\frac{(x^2-4)-2}{(x^2-4)+3} = \frac{x^2-6}{x^2-1}$ on the restriction of $1 \leq x < 2$.

$$x^2 - 4 \geq 0 \quad \rightarrow \quad x \leq -2 \ \ or \ \ x \geq 2$$

The intersection of $x \leq -2 \ or \ x \geq 2$ and $x \geq 1$ is $x \geq 2$. Giving the composed expression

$3(x^2 - 4) + 1 = 3x^2 - 11$ on the restriction of $x \geq 2$.

Therefore

$$(f \circ g)(x) = \begin{cases} \sqrt{x+5}, & -5 \leq x < 1 \\ \dfrac{x^2-6}{x^2-1}, & 1 \leq x < 2 \\ 3x^2 - 11, & x \geq 2 \end{cases}$$

Notice the domain of the function can be calculated by determining the domain of each expression on their corresponding restrictions.

$$\sqrt{x+5} \quad \rightarrow \quad x + 5 \geq 0 \quad \rightarrow \quad x \geq -5$$

This is only for the restriction of $-5 \le x < 1$.

$$\frac{x^2 - 6}{x^2 - 1} \quad \rightarrow \quad x^2 - 1 \ne 0 \quad \rightarrow \quad x \ne \pm 1$$

This is only for the restriction of $1 \le x < 2$, so $x \ne -1$ is the only value excluded.

$$3x^2 - 11$$

The domain of this expression is all real numbers.

Therefore

$$dom(f \circ g) = [-5,1) \cup (1, \infty).$$

Exercises:

1. Given the piecewise functions

$$f(x) = \begin{cases} -5x + 1, & x < -2 \\ x - 4, & x \ge -2 \end{cases}$$

and

$$g(x) = \begin{cases} 3x - 2, & x < 1 \\ -2x + 3, & x \ge 1 \end{cases}$$

   a. Determine the value of $f(g(-3))$.
   b. Determine the value of $f(g(4))$.
   c. Determine the function expression for the composition $f(g(x))$.
   d. Use the function expression found in c. to verify exercises a. and b.

2. Determine the domain of the composition $f(g(x))$ given the piecewise functions

$$f(x) = \begin{cases} -5x + 1, & x < 0 \\ \sqrt{x - 1}, & x \ge 0 \end{cases}$$

and

$$g(x) = \begin{cases} \dfrac{x - 5}{x - 4}, & x < 2 \\ x^2 + 7, & x \ge 2 \end{cases}$$

3. Determine the function expression for the piecewise function in exercise 2. Use the function expression to verify the domain.

The Arithmetic Derivative

Paul Bedard

Introduction

One of the best tricks I know for understanding a mathematical concept more fully is to apply the concept inappropriately. That is, to use a rule or an algorithm in a situation never intended for its use. Whoever first put chocolate on a pretzel understood this. Many of the great mathematical discoveries have come about through the creative misapplication of ideas. Indeed, without a great deal of flexibility in the matter of what we permit ourselves to think about, modern mathematics as we know it wouldn't exist. (WE must not forget, however, that some nonsense is truly and simply nonsense. When playing in the sandbox, the refusal to cling to traditional architectural forms must be accompanied by a willing fist, ready to quickly obliterate grotesqueries.)

Against this notion, Simon Stevin said of complex numbers in 1585, "There is enough legitimate matter, even infinitely much, to exercise oneself without occupying oneself and wasting time on uncertainties." (Swetz, p. 138). But Leibniz, also discussing the same subject, said "One would never have believed that $\sqrt{1 + \sqrt{-3}} + \sqrt{1 - \sqrt{-3}}$ make $\sqrt{6}$ and there is something hidden therein which is incomprehensible to me." (Swetz, p. 138).

Finding such hidden somethings in apparently nonsensical or meaningless applications (such as applying the square root concept to negative numbers) is a fruitful source of new mathematics. But it is also a way to get a better grasp of the old.

In this essay I will show how the product rule for derivatives can be applied to integers, creating a sequence of integers referred to as "arithmetic derivatives." If you have no familiarity with calculus, please read on nonetheless. Though the first part of the essay, in which I review the product rule as used in calculus, might not mean much, the second part can be understood readily by anyone who can multiply and add whole numbers. You might achieve an appreciation of a calculus technique without ever knowing what calculus is!

The Leibniz Rule

The Leibniz rule, often called the product rule, is a rule intended to facilitate the differentiation of products of differentiable functions. The rule, stated in Leibniz or differential notation, is

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$$

Here u and v are differentiable functions of x. The right side of the definition may be read:

"The first times the derivative of the second, plus the second times the derivative of the first."

(I can't neglect to point out that the most common student error when taking derivatives of products, what I call the "first false product rule," – the false idea that the derivative of a product is the product of the derivatives – was an error first made by Leibniz himself. However, as I tell my students often, making errors no matter how silly is not a problem. The problem is failing to recognize and correct them.)

Example 1. A calculus derivative with and without the Leibniz rule.

Find the derivative of $y = (2x + 1)(3x - 1)$

We may do this in either of two ways. Multiplying the binomials before differentiating reduces the problem to

$$\frac{d}{dx}(6x^2 + x - 1)$$

Applying the power rule (and the sum rule) quickly yields

$$\frac{d}{dx}(6x^2 + x - 1) = 12x + 1$$

We don't need the product rule at all. But it is instructive to see how it works here.

$$\frac{d}{dx}(2x + 1)(3x - 1) = (2x + 1)\frac{d}{dx}(3x - 1) + (3x - 1)\frac{d}{dx}(2x + 1)$$

Now we can apply the power rule to find each of the derivatives and this leads to

$$(2x + 1) \cdot 3 + (3x - 1) \cdot 2 = 6x + 3 + 6x - 2 = 12x + 1$$

There is nothing that builds the confidence in a new rule as much as seeing it yield the same answer as a trusted old rule!
The power rule, which we will see again in a new context later, applies to derivatives of power functions of x.

The Power Rule for (Calculus) Derivatives

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

The Arithmetic Derivative

"The **arithmetic derivative** n' was introduced, probably for the first time, by the Spanish mathematician José Mingot Shelly in June 1911 with "Una cuestión de la teoría de los números", work presented at the "Tercer Congreso Nacional para el Progreso de las Ciencias, Granada."
(From The Online Encyclopedia of Integer Sequences, https://oeis.org/A003415).
The calculus derivative of any constant function is identically equal to zero. This means that the derivative of a simple number, say 5, when thought of as a function $f(x) = 5$, is zero. The arithmetic derivative is not a calculus derivative. We are going to define a new function on the integers and call it an arithmetic derivative simply because it reproduces some properties of calculus derivatives. In fact, we are going to define this function in terms of one of these properties – the Leibniz rule.
We want the arithmetic derivative of a positive integer to be another positive integer, and we want the Leibniz rule to be satisfied.

Definition. The <u>**arithmetic derivative**</u> of a positive integer N is a unique positive integer N' such that
     (1) If N is prime then $N' = 1$,
     (2) If $N = a \cdot b$ for any positive integers a and b then $N' = ab' + ba'$

(N' may be read as "N prime"; prime notation here is borrowed from calculus.)
There are some things we must note about this definition. First, it doesn't directly construct arithmetic derivatives, except in the case of primes. In other words, it provides no formula for calculating an arithmetic derivative. Yet, it is a perfectly acceptable definition. We will examine why this is so.

First, we should ask ourselves whether N', defined this way, will be unique. In other words, is this a function (just one output for each input) on the positive integers at all? If an integer can be shown to have two different arithmetic derivatives, both of which satisfy the definition, then we have a problem. And in fact, it is reasonable to be concerned. Notice we said "if $N = a \cdot b$ for any positive integers a and b". Obviously, many integers can be factored into a product of two other integers in many ways.

Example 2. Multiplicity of representations.

$$120 = 1 \cdot 120 = 2 \cdot 60 = 3 \cdot 40 = 4 \cdot 30 = 5 \cdot 24 = 6 \cdot 20 = 8 \cdot 15 = 10 \cdot 12$$

In order for our definition to yield a unique arithmetic derivative for 120, the values of the arithmetic derivatives of all of these factors must be such that no matter which factoring we choose, when we apply the Leibniz rule, we get the same answer.
Let's start by stating a theorem.

Theorem 1. $1' = 0$

Proof. Let N be any positive integer. Then $N = N \cdot 1$ and therefore $N' = N \cdot 1' + 1 \cdot N'$. Subtracting N' from both sides yields $0 = N \cdot 1'$. By the zero product property of multiplication (which says that, if a product equals zero, at least one of the factor must equal zero), since $N \neq 0$ then $1' = 0$.
<div align="center">■</div>

When we find a general rule for calculating arithmetic derivatives of any positive integer, we will see that the arithmetic derivative of a number is the same no matter how we factor the number.

Calculating Arithmetic Derivatives

Let's begin to create a table of arithmetic derivatives. It will be very incomplete at this point.

Table 1 (a)

N	N'		N	N'
1	0		9	
2	1		10	
3	1		11	1
4			12	
5	1		13	1
6			14	
7	1		15	
8			16	

We have filled in the arithmetic derivative of 1, from theorem 1, and the arithmetic derivatives of the primes listed, which are equal to 1 by definition.

　　　　Can we find the arithmetic derivative of a composite number? If the composite number is a product of two primes, we surely can. Let's find $6'$.

Example 3. The arithmetic derivative of a product of two primes.

$$6' = (2 \cdot 3)' = 2 \cdot 3' + 3 \cdot 2' = 2 \cdot 1 + 3 \cdot 1 = 5$$

Can we come up with any more general formula from this example? We can!

Theorem 2. If $N = p \cdot q$ where p and q are primes, then $N' = p + q$.

Proof.   Let $N = p \cdot q$ where p and q are primes.
　　　　Then $N' = (p \cdot q)' = p \cdot q' + q \cdot p' = p \cdot 1 + q \cdot 1 = p + q$　　　　　　　　　■

Don't worry – the proofs will get less trivial.

From theorem 2, we can determine the arithmetic derivatives of 10, 14, and 15.

Example 4. **More arithmetic derivatives of products of two primes.**

$$10' = 2 + 5 = 7, \ 14' = 2 + 7 = 9, \ 15' = 5 + 3 = 8$$

Let's add these to our table.

Table 1 (b)

N	N'		N	N'
1	0		9	
2	1		10	7
3	1		11	1
4			12	
5	1		13	1
6	5		14	9
7	1		15	8
8			16	

So far, we have a rather nifty idea – a function which assigns to composite numbers which can be factored as the product of two primes, the sum of those primes. But what about a number which is a product of two nondistinct primes – for instance, 9 which is the product of 3 and 3. The theorem should work for such numbers also.

Example 5. The arithmetic derivative of a prime squared.

$$9' = (3 \cdot 3)' = 3 + 3 = 6$$

But what about number which are the product of more than two primes? Like 8 which is 2^3 (the product of three primes, all the same) and 12 which is $2^2 \cdot 3$? And what about 16 which is 2^4?

The next theorem will help us with 8 and 16. We are slowly building to a general formula, never fear.

Theorem 3. If $N = p^n$ where p, n are positive integers and p is prime, then $N' = np^{n-1}$.

Proof. Let $N = p^n$ where p, n are positive integers and p is prime.

If $n = 1$, then the theorem holds since $N' = p' = 1$ by definition, and by the theorem $(p^1)' = 1 \cdot p^{1-1} = 1 \cdot p^0 = 1$.

We will proceed by induction. Suppose the theorem holds for $n = k - 1$. Then for $N = p^{k-1}, N' = (k-1)p^{k-2}$.

Suppose $N = p^k$. Then $N' = (p^k)' = (p \cdot p^{k-1})' = p \cdot (p^{k-1})' + (p^{k-1}) \cdot p'$

Since $(p^{k-1})' = (k-1)p^{k-2}$ and $p' = 1$, we get

$N' = p \cdot (k-1)p^{k-2} + (p^{k-1}) \cdot 1 = kp^{k-1}$. ∎

Did you notice anything special about theorem 3? It looks a lot like the power rule for calculus derivatives! Let's take a moment to reflect on what has happened here. We started by defining a new kind of derivative, not based at all on calculus derivatives except that we insisted on the Leibniz rule. There was no particular reason to expect arithmetic derivatives to share anything else in common with calculus derivatives. But we now see that the power rule has come along for the ride!

What does this tell us about the relationship between the power rule and the Leibniz rule? Already, our outside-the-cube thinking has had positive results. We are tricking ourselves into examining relationships within "real" mathematics more closely!

Example 6. Arithmetic derivatives of primes to any power.

$$4' = (2^2)' = 2 \cdot 2^1 = 4; \quad 16' = (2^4)' = 4 \cdot 2^3 = 32.$$

Exercise 1. Find $8'$ both by factoring 8 into 2×4 and by applying theorem 3.

Now we can fill in the table a little more.

Table 1 (c)

N	N'		N	N'
1	0		9	6
2	1		10	7
3	1		11	1
4	4		12	
5	1		13	1
6	5		14	9
7	1		15	8
8	12		16	32

A General Formula for Arithmetic Derivatives (I)

Our next task is a bit more daunting. We would like to create a formula which will give us the arithmetic derivative of any positive integer, not just a power of a prime or a product of two primes.

To accomplish this, we need two further facts. One is the extended Leibniz rule for calculus derivatives. Suppose we have a function which is a product of an arbitrary number of differentiable functions.

Extended Leibniz Rule for calculus derivatives.

$$\frac{d}{dx}(u_1 u_2 u_3 \ldots u_k) = \left(\frac{d}{dx}u_1\right)u_2 u_3 \ldots u_k + u_1\left(\frac{d}{dx}u_2\right)u_3 \ldots u_k + \cdots + u_1 u_2 u_3 \ldots \left(\frac{d}{dx}u_k\right)$$

To understand what is going on in this rule, look carefully at each term. Each term is the product of all but one of the functions and the derivative of that missing function. The derivative can be thought of as moving like a wave through the products.

We present the extended Leibniz rule without proof, but we will prove the arithmetic derivative version and this will give a solid hint for a proof of the calculus version.

Theorem 4. Extended Leibniz Rule for Arithmetic Derivatives

Let $N = a_1 a_2 a_3 \ldots a_n$, for positive integers a_i. Then

$$N' = (a_1)'a_2 a_3 \ldots a_n + a_1(a_2)'a_3 \ldots a_n + a_1 a_2(a_3)' \ldots a_n + \ldots + a_1 a_2 a_3 \ldots (a_n)'$$

Proof. Again, we will use induction. For $n = 1$, we have $N' = (a_1)'$ and the theorem is satisfied.

Assume the theorem is true for $n = k - 1$. Then for $M = a_1 a_2 a_3 \ldots a_{k-1}$,
$$M' = (a_1)'a_2 a_3 \ldots a_{k-1} + a_1(a_2)'a_3 \ldots a_{k-1} + \ldots + a_1 a_2 a_3 \ldots (a_{k-1})'. \quad (*)$$

Now for any positive integer a_k, let $N = a_1 a_2 a_3 \ldots a_{k-1}a_k$.
Then $N' = (a_1 a_2 a_3 \ldots a_{k-1}a_k)'$
But by the definition,
$(a_1 a_2 a_3 \ldots a_{k-1}a_k)' = a_1 a_2 a_3 \ldots a_{k-1} \cdot (a_k)' + a_k \cdot (a_1 a_2 a_3 \ldots a_{k-1})'$
Substituting (*) we get
$(a_1 a_2 a_3 \ldots a_{k-1}a_k)' =$
$a_1 a_2 a_3 \ldots a_{k-1} \cdot (a_k)' + a_k \cdot ((a_1)'a_2 a_3 \ldots a_{k-1} + \qquad a_1(a_2)'a_3 \ldots a_{k-1} + \ldots + a_1 a_2 a_3 \ldots (a_{k-1})')$
Distributing a_k gives us
$N' = (a_1)'a_2 a_3 \ldots a_k + a_1(a_2)'a_3 \ldots a_k + \ldots + a_1 a_2 a_3 \ldots (a_k)'.$ ∎

Exercise 2. Find the arithmetic derivative of $210 = 2 \cdot 3 \cdot 5 \cdot 7$

Combining this extended rule with our previous rules will complete our task of finding a general formula for the arithmetic derivative of any positive integer. Before we state it we need to recall the fundamental theorem of arithmetic. Every positive integer has a unique prime factorization.

The Fundamental Theorem of Arithmetic

> For any positive integer N there exists a unique factorization $N = p_1{}^{n_1} p_2{}^{n_2} p_3{}^{n_3} \dots p_k{}^{n_k}$
> where p_i, n_i are positive integers and p_i is prime.

For example, $12 = 2^2 \cdot 3$ and $72 = 2^3 \cdot 3^2$. The order of the primes does not matter (sometimes we say "every positive integer has a unique prime factorization, up to order" to be clearer.) It is customary to list the primes in increasing order. If 1 were considered a prime number (it's not) the fundamental theorem would be invalid since we could select any power for the 1 and the factorization would not be unique.

A Combinatorial Digression

Before we present our main theorem, the formula for the general arithmetic derivative, I want to point out that theorem 4 links the topic of arithmetic derivatives with the topic of combinatorics, which, loosely defined, is the branch of mathematics that deals with counting finite collections of objects.

Example 7. Selecting a committee.

Given ten people, how many ways are there to select a set of 6 of them? The order of the people selected does not matter – suppose that you are forming a committee.

The answer to this question introduces a concept called combinations. We write $\binom{10}{6}$

and read this as "10 choose 6." The formula for evaluating this expression involves factorials. Factorials are written with an "!"; for example, $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$.

Formula for calculating number of combinations when k objects are selected from a set of n objects, $k \le n$.

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

So the solution to example 7 is

$$\binom{10}{6} = \frac{10!}{6!\,(10-6)!} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6!}{6! \cdot 4!} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

How does this tie in with arithmetic derivatives? The next example should clarify this.

Example 8. Arithmetic derivatives and combinations.

> Find the arithmetic derivative of $2 \cdot 3 \cdot 5$
> Using theorem 4, we see that $(2 \cdot 3 \cdot 5)' = 2' \cdot 3 \cdot 5 + 2 \cdot 3' \cdot 5 + 2 \cdot 3 \cdot 5'$
> But since 2,3,5 are all prime and thus have arithmetic derivatives of 1, this simplifies to $(2 \cdot 3 \cdot 5)' = 3 \cdot 5 + 2 \cdot 5 + 2 \cdot 3$.

Look at this result carefully. Note that the arithmetic derivative of a product of three prime factors turned out to be every possible product of two of those factors. There turned out to be th5ree such products. Counting the products is a matter of using the combinatorial formula for combinations.

$$\binom{3}{2} = 3$$

Recall the exercise above, when you found $210'$. $210 = 2 \cdot 3 \cdot 5 \cdot 7$. We needed to find every possible product of three of those factors.

$$\binom{4}{3} = 4$$

For any integer which is a product of n distinct prime factors, the arithmetic derivative will be the sum of $\binom{n}{n-1}$ products.

Example 6. **Number of distinct products.**

$$60' = (2 \cdot 2 \cdot 3 \cdot 5)' = 2' \cdot 2 \cdot 3 \cdot 5 + 2 \cdot 2' \cdot 3 \cdot 5 + 2 \cdot 2 \cdot 3' \cdot 5 + 2 \cdot 2 \cdot 3 \cdot 5'$$

This is equal to $2(2 \cdot 3 \cdot 5) + 2 \cdot 2 \cdot 5 + 2 \cdot 2 \cdot 3$

There are $\binom{4}{3}$ products but only 3 distinct products. How can we count the distinct products?

Exercise 3. In calculating the arithmetic derivative of $3^4 \cdot 5^2 \cdot 11^3$, how many products would there be? How many distinct products would there be?

A General Formula for Arithmetic Derivatives (II)

Now let's return to our main theorem. We have seen how the extended Leibniz rule allows us to take an arithmetic derivative of a product with an arbitrary number of factors. We have also seen that every positive integer can be written as a unique product of primes to powers.

Theorem 5. General formula for arithmetic derivatives.
Let N be a positive integer with prime factorization $p_1{}^{n_1} p_2{}^{n_2} p_3{}^{n_3} \dots p_k{}^{n_k}$. Then
$N' = N(n_1 p_1{}^{-1} + n_2 p_2{}^{-1} + \dots + n_k p_k{}^{-1})$

Proof.
$N' = (p_1{}^{n_1})' p_2{}^{n_2} p_3{}^{n_3} \dots p_k{}^{n_k} + p_1{}^{n_1} (p_2{}^{n_2})' p_3{}^{n_3} \dots p_k{}^{n_k} + \dots + p_1{}^{n_1} p_2{}^{n_2} p_3{}^{n_3} \dots (p_k{}^{n_k})'$
$= n_1 p_1{}^{n_1-1} p_2{}^{n_2} p_3{}^{n_3} \dots p_k{}^{n_k} + p_1{}^{n_1} n_2 p_2{}^{n_2-1} p_3{}^{n_3} \dots p_k{}^{n_k} + \dots + p_1{}^{n_1} p_2{}^{n_2} p_3{}^{n_3} \dots n_k p_k{}^{k-1}$
We factor out $p_1{}^{n_1} p_2{}^{n_2} p_3{}^{n_3} \dots p_k{}^{n_k}$ which gives us $N(n_1 p_1{}^{-1} + n_2 p_2{}^{-1} + \dots + n_k p_k{}^{-1})$ ∎

Example 7 Find $12'$.

$$12' = (2^2 \cdot 3)' = 12\left(\frac{2}{2} + \frac{1}{3}\right) = 12\left(\frac{4}{3}\right) = 16$$

Finally, we can complete our table of the first 16 arithmetic derivatives:

Table 1 (d)

N	N'		N	N'
1	0		9	6
2	1		10	7
3	1		11	1
4	4		12	16
5	**1**		13	1
6	**5**		14	9
7	1		15	8
8	12		16	32

Exercise 5. Find the arithmetic derivative of $2 \cdot 3^4 \cdot 5 \cdot 7^2 = 39{,}690$

Second Arithmetic Derivatives

We will define a second arithmetic derivative just as we do with calculus derivatives – it will be the arithmetic derivative of an arithmetic derivative, and will be denoted N'', which we can read as "N double prime."

We can partially fill in a table of second derivatives by taking the entry in the N' column and looking for that number in the N column. For instance, consider $N = 6$ as noted above in boldface. $6' = 5$, so we find 5 in the N column and note that $5' = 1$. So $6'' = 1$.

By doing this as often as we can, let's begin a table of first and second arithmetic derivatives.

Table 2

N	N'	N''		N	N'	N''
1	0			9	6	5
2	1	0		10	7	1
3	1	0		11	1	0
4	4	4		12	16	32
5	**1**	0		13	1	0
6	**5**	1		14	9	6
7	1	0		15	8	12
8	12	16		16	32	

This was pretty successful! Note that $1''$ currently lacks any defined answer, since we have defined arithmetic derivatives as a function on positive integers. We will also feel this lack if we try to find third or higher order arithmetic derivatives, since for any prime $p, p'' = 0$. We had better follow convention and define $0' = 0$.

Let's find more comparisons between calculus and arithmetic derivatives. A power function is said to have (calculus) derivatives that "vanish," because for sufficiently high order, the derivative of that order is zero, and since the calculus derivative of zero is zero, any derivative of higher order is zero as well.

Example 8. A function whose calculus derivative vanishes.

$$f(x) = x^3, f'(x) = 3x^2, f''(x) = 6x, f'''(x) = 6, f^{(4)}(x) = 0$$

Since the derivative of 0 is 0, every derivative after the fourth is also equal to zero. So we can state:

$$\frac{d^n}{dx^n}(x^3) = 0, n \geq 4$$

We can see from the table that arithmetic derivatives may vanish also. In fact, the arithmetic derivatives of primes vanish at the second derivative. $10'' = 1$, so the arithmetic derivative of 10 vanishes at the third derivative.

Arithmetic Differential Equations

In calculus, an equation such as $y' = 2x$ can be solved by finding a function whose derivative is $2x$. This is called a differential equation because it contains derivatives. Differential equations can become much more complicated: $y'' + 2y' + y = 0$ is one example. Students of calculus learn a large variety of methods for finding functions which are solutions to such equations.

General methods for solving arithmetic differential equations would be very different from calculus methods. The analogy breaks down badly here. I am not aware of any work that has been done on such methods. However, it is interesting and instructive to look at some simple examples.

Example 9. **An arithmetic differential equation.**

Solve the equation $x' = 14$.

We have seen that numbers which are the product of two distinct primes have arithmetic derivative which is the sum of those primes. (Theorem 2). Let's look at all the ways that 14 can be written as a sum of positive integers.

$$14 = 1 + 13 = 2 + 12 = 3 + 11 = 4 + 10 = 5 + 9 = 6 + 8 = 7 + 7.$$

3 and 11 are primes, so their product 33 has arithmetic derivative $3 + 11 = 14$. 7 is also a prime so 49 is another solution. We have no reason to conclude that we have found all solutions. Unlike the situation with calculus, finding some solutions doesn't make finding further solutions any easier.

Example 10. **Another differential equation.**

Solve $x' = x$

In calculus, there is a set of functions which are their own derivatives – multiples of the exponential function $f(x) = e^x$.
Examining table 2, we can see that 4 is its own arithmetic derivative. That is the only solution on our table.
However, perhaps we can find a pattern and generate some more solutions.
$4 = 2^2$. Recall theore 3, the "power rule" for arithmetic derivatives. Now consider any integer of the form p^p, p prime. Its arithmetic derivative would be $p \cdot p^{p-1} = p^p$.
Thus, all integers of the form p^p, p prime are solutions to $x' = x$.
24 and 256 are additional examples.

I invite readers to propose and solve further arithmetic differential equations. We can always reverse engineer them from the answers in the table (for example, since $3' = 1$, we can create $N = 3N'$. Indeed, p is a solution to $N = pN'$ for any prime p.) Try to do better.

Some Differences Between Calculus and Arithmetic Derivatives

We have seen that arithmetic derivatives, which obey the Leibniz rule by definition, also have a version of the power rule. They can be extended to derivatives of higher order, and they can vanish or remain constant, just like calculus derivatives. However, they differ in almost every way. We must not confuse our use of the term "derivative" for arithmetic derivatives to convey any more than a superficial, if pleasing and sometimes surprising, comparison.
One of the most helpful and basic rules for calculus derivatives is the sum rule:

$$\frac{d}{dx}(f + g) = \frac{df}{dx} + \frac{dg}{dx}$$

This rule does not hold for arithmetic derivatives. Note that $30' = 31 \neq 25' + 5'$.
Another difference is more fundamental. Calculus derivatives give us formulas for slopes of tangent lines to curves. We have not discovered any useful facts about numbers that their arithmetic derivatives provide. Not yet! Perhaps further investigation of this intriguing if highly artificial concept may yet lead to some good mathematics. However, as I hope I have shown, even if it doesn't, the investigation is worthwhile because it helps us think logically, it helps us focus on the properties of derivatives, and it is intrinsically intriguing.

Postscript: An Alternative Arithmetic Derivative

In the definition of arithmetic derivative, we specified that $p' = 1$ for any prime. We could not prove this if we had not stated it. It is not implied by the Leibniz rule, but it was convenient for us.
Suppose we had said, for instance, $p' = 2$ for any prime. This would give us another arithmetic derivative altogether. Let's see if it would work. We will use the notation $N^{\sim}$ for arithmetic derivatives of this type. We need to see if $N^{\sim}$ is unique for every N.

Example 11. Uniqueness of an alternate derivative.

$12 = 2 \cdot 6 = 3 \cdot 4$

$6\tilde{} = (2 \cdot 3)\tilde{} = 2 \cdot 3\tilde{} + 3 \cdot 2\tilde{} = 2 \cdot 2 + 3 \cdot 2 = 10$
$4\tilde{} = (2 \cdot 2)\tilde{} = 2 \cdot 2\tilde{} + 2 \cdot 2\tilde{} = 2 \cdot 2 + 2 \cdot 2 = 8$

$12\tilde{} = (2 \cdot 6)\tilde{} = 2 \cdot 6\tilde{} + 6 \cdot 2\tilde{} = 2 \cdot 10 + 6 \cdot 2 = 32$
$12\tilde{} = (3 \cdot 4)\tilde{} = 3 \cdot 4\tilde{} + 4 \cdot 3\tilde{} = 3 \cdot 8 + 4 \cdot 2 = 32$

What if we defined a derivative, say N^*, such that $p^* = p$ for any prime p. Would this work? Let's let this be your final exercise.

http://math.arizona.edu/~ura-reports/063/Sandhu.Aliana/Final.pdf

Modeling Biological Processes with Discrete Time Markov Chains (DTMC)

Janice Fritz

Introduction

I plan to investigate the use and limitations of Discrete Time Markov Chains (DTMC) in modeling biological processes. DTCM depend on stochastic matrices and they are used to model processes in which the future depends only on the current state, not on the past. I've found several examples of DTMC-based models in Biology, including cell proliferation, cell migration, birth-death processes, enzyme kinetics, and DNA sequences. I hope to design my own simple model of a biological process (possibly cell differentiation, but I'm not willing to commit to that at this point) with a stochastic matrix that I can use to predict the result of the process. This is what I would demonstrate on my mathematics page.

Modeling Biological Processes with Discrete Time Markov Chains (DTMC)

Markov processes are named after Russian mathematician Andrey Markov (1856-1922) [1]. Markov, whose younger brother and son were also notable mathematicians, worked on stochastic processes [1]. Stochastic processes involve random variables that change over time (or some other parameter) based on some probability [2, 4]. A Markov process is a type of stochastic process in which the outcome at any point depends only on the previous state and not on how the system arrived at that state [3]. It also requires that the number of states is finite and probabilities are constant [4].

As an example of a situation obeying the rules of a Markov process, consider the hypothetical intersection of two one-way streets, Elm St. and North St., at which 45% of vehicles approaching on Elm St. go straight and 55% turn onto North St.. Of the vehicles approaching the intersection from North St., 75% go straight and 25% turn onto Elm St.. In this situation, the probability of a car continuing on its current street or turning onto the other street depends only on which street it approaches the intersection, not where the car was before approaching the intersection. Whether you are coming from the library or from the grocery store, if you are driving on North St., you are more likely to stay on North St. than to turn onto Elm.

The outcome of a Markov process can be determined using a stochastic matrix. See Example 1 in the mathematics section to see how to determine the number of cars that will end up on North and Elm streets after the intersection.

For some processes, the stochastic matrix can be applied to the outcome repeatedly to model the results over many cycles of the process. While the simple example above does not work this way since a vehicle won't encounter the intersection again after passing it, we can devise a more general example that does.

Instead of a single intersection, consider a grid of North/South (N/S) and East/West (E/W) streets in which 15% of vehicles traveling on a N/S street turn onto the E/W street at any given intersection and 85% go straight to remain on the N/S street. Of the cars traveling on a E/W street, 5% will turn onto the N/S street at any given intersection and 95% go straight. Once a car goes through an intersection, it will eventually reach another intersection where it will turn or go straight based on the probabilities appropriate for the direction it is now going. See Example 2 in the mathematics section for an example.

If a Markov process occurs at discrete time intervals or between discrete states rather than continually over time, it is a discrete time Markov chain (DTMC). Many Biological processes fit this model of changes over discrete time or between discrete states.

At the population level, DTMCs can predict changes in population size and the time until the extinction of a population [5]. This model is based on the probabilities of a birth increasing the population by one, a death decreasing the population by one, or no change occurring during a given time interval. In order for the change in population to be limited to one, that is, not more than one birth or death occurring during the time interval, the time interval has to be sufficiently short [5]. This model is not as simple as those above because the probabilities of birth or death change as the number of individuals in the population changes, so the stochastic matrix changes with each iteration. Over many time intervals, the population will grow indefinitely, become extinct or reach a stable sustainable level.

An example of using DTMCs at the molecular level is the modeling of changes in protein conformation [4]. Many protein molecules experience changes in their shape and this change in shape can affect how they function. Consider a protein that has three conformational states, C_1, C_2, and C_3. Each second, there is a certain probability that a molecule of the protein will switch to one of the other shapes. These probabilities can be used to create a transition matrix to predict the steady state amount of protein in each conformation [4]. This is of biological significance because the different conformations may have different strength, different binding affinity or different efficiency in catalyzing reactions. Changes in cellular conditions, such as pH, temperature, or presence of other compounds may affect the probabilities of switching between conformations, leading to a more or less effective population of protein molecules.

I am interested in modeling cellular differentiation using DTMCs to predict the proportion of different cell types that would be present after a certain length of time. Cellular differentiation is the process of a stem cell becoming a specific mature cell type [6]. All the different types of cells in the human body are the descendants of a single zygote formed when a sperm fertilized an egg. The process begins with a mass of totipotent cells formed by repeated cell division of the zygote. Totipotent cells are not differentiated at all. They have the ability to give rise to any type of cell. In the formation of an embryo, the totipotent cells differentiate into one of three cell types: endoderm, mesoderm, and ectoderm. These cells are called pluripotent because they can give rise many different cell types, but not all. Endoderm cells eventually differentiate in to muscle, lung, and gut cells; mesoderm cells give rise to connective tissues like cartilage, bone, blood, and fat; and ectoderm cells become skin cells and neurons [6, 7]. This process involves many steps though multipotent cells types that have been left out for simplicity.

The standard view of cellular differentiation is that it is a one-way process. Once a cell has differentiated, it may differentiate further along the same path into an even more specialized cell, but it cannot revert back to a less differentiated cell type and take a different path. See the figure below from [6].
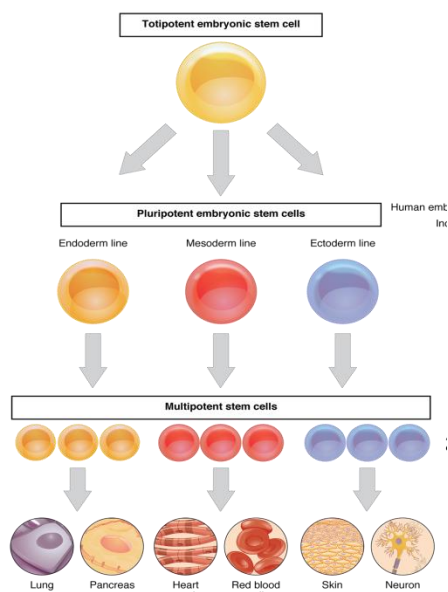
However, there are situations in which reversion and re-differentiation may occur. It is necessary in animals that regrow amputated body parts, such as amphibians and lizards growing a new tail. There is some evidence that mammalian cells may also dedifferentiate under certain conditions [8]. An important area of current research is the use of induced pluripotent stem cells. These are differentiated adult cells that are artificially converted back into stem cells that can become many different cell types to create tissues and organs ex vivo [6].

While the standard process would still be interesting to model, I chose to create a DTMC model of cellular differentiation that included the possibility of some dedifferentiation, although this would be a rare event. A Spanish Wikipedia page provided the visual model of the process [9]:
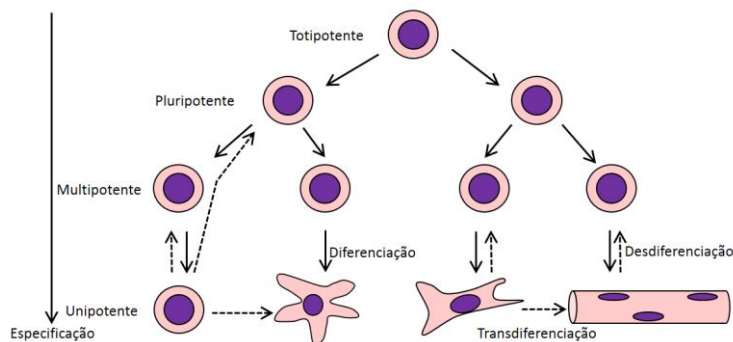
The first thing to consider in modeling the differentiation shown in the figure above is the transition probabilities. This model is extremely simplified in both the number of cell types and the number of steps. For example, a population of totipotent cells differentiates into more than two types of pluripotent cells, which

divide into more than two types of multipotent cells, which go through more than one step to become fully differentiated adult cells. Also, since the figure itself is only a representation of the general process instead of specific known pathways of differentiation, the transition probabilities are completely invented. Except for the totipotent cells, which all differentiate into pluripotent cells, there is at least a small probability that a cell type will remain the same rather than differentiating to the next level. The probability of dedifferentiating (moving back to a less differentiated cell type) is very small. See example 3 on the mathematics section for the 11x11 transition matrix.

The transition matrix has some interesting properties I noticed as I was setting it up. As with any transition matrix, the main diagonal indicates the probability of remaining in the same state. A quick glance at these values shows with cell types are transitory steps and which terminally differentiated cells that are unlikely to change. One cells type that does not follow this pattern is the cell type labelled C2. The represents a cell type like the spermatogonia in the testes. Every time a spermatogonium divides, it produces one spermatogonium and one spermatocyte. The spermatocyte will give rise to four sperm while the new spermatogonium with divide to produce a spermatocyte and another spermatogonium. This continual self-renewing of the spermatogonia allows men to produce sperm throughout their life time instead of eventually running out. Another interesting feature of the transition matrix is that values below the main diagonal are the probabilities of transition to a more differentiated state while values above the main diagonal are the probabilities of dedifferentiation. The later values are fewer and smaller than the former.

In order to see how a population of cells will change over time, calculate the product of the transition matrix and a matrix representing the original cell population. To represent the original population of undifferentiated cells, we will assume there are 100 totipotent cells and no other cells types, a 11x1 matrix with 100 as the initial entry and zeros in the other positions. Multiplying this by the transition matrix yields a population with no totipotent cells and two types of pluripotent cells. Multiplying by the transition matrix again produces the next generation of cells. Repeating many times shows the eventual population of differentiated cells after n generations: $P_n = T^n P_o$. See example 3 in the mathematics section.

In this model, the discrete time interval for the Markov chain is one cell cycle. Since the population of cells will double with each time interval, the population matrix must be multiplied by 2 for each time interval or 2^n for each time interval. Since the population matrix changes each interval but the transition matrix remains the same, we can multiply the transition matrix by the scalar 2 to get $P_n = (2T)^n P_o$, the population of cells, instead of just the proportion of cells, after n generations.

Mathematics

Example 1. The intersection of North and Elm*

A stochastic matrix represents probabilities staying on the same street or turning onto the other:

Starting on Elm	Starting on North	
0.45	0.25	Ending on Elm
0.55	0.75	Ending on North

The product of this matrix and a matrix representing the number of cars approaching the intersection on each street will predict the number of cars on each street after the intersection. If 40 cars approach the intersection on Elm, and 160 cars approach the intersection on North,

$$\begin{bmatrix} 0.45 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \times \begin{bmatrix} 40 \\ 160 \end{bmatrix} = \begin{bmatrix} 58 \\ 148 \end{bmatrix}$$

there will be 58 cars on Elm and 148 cars on North on the other side of the intersection.

Example 2. The intersection of North/South and East/West streets

Stochastic matrix representing the probability staying on the same street or turning onto the other:

Starting on N/S	Starting on E/W	
0.95	0.15	Ending on N/S
0.05	0.85	Ending on E/W

If there are 340 cars on N/S streets and 80 cars on E/W streets, after these cars reach an intersection,

$$\begin{bmatrix} 0.95 & 0.15 \\ 0.05 & 0.85 \end{bmatrix} \times \begin{bmatrix} 340 \\ 80 \end{bmatrix} = \begin{bmatrix} 335 \\ 85 \end{bmatrix}$$

there will be 335 cars on N/S streets and 85 cars on E/W streets. After these cars reach the next intersection on their respective routes,

$$\begin{bmatrix} 0.85 & 0.15 \\ 0.05 & 0.95 \end{bmatrix} \times \begin{bmatrix} 335 \\ 85 \end{bmatrix} = \begin{bmatrix} 331 \\ 89 \end{bmatrix}$$

there will be 331 cars on N/S streets and 89 cars on E/W streets. If this process continues through many intersections, the number of cars on the N/S and E/W streets reaches a steady state where the number of cars turning onto E/W streets is the same as the number of cars turning onto N/S streets to the outcome will no longer change.

Example 3. Cellular Differentiation

Cell differentiation model:
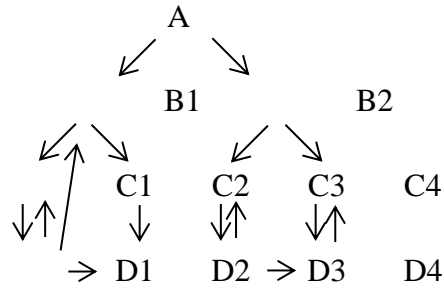
Transition matrix [T]:

From	A	B1	B2	C1	C2	C3	C4	D1	D2	D3	D4	To
	0	0	0	0	0	0	0	0	0	0	0	A
	0.6	0.1	0	0	0	0	0	0.005	0	0	0	B1
	0.4	0	0.05	0	0	0	0	0	0	0	0	B2
	0	0.70	0	0.05	0	0	0	0.01	0	0	0	C1
	0	0.2	0	0	0.5	0	0	0	0	0	0	C2
	0	0	0.45	0	0	0.05	0	0	0	0.005	0	C3
	0	0	0.5	0	0	0	0.2	0	0	0	0.001	C4
	0	0	0	0.95	0	0	0	0.95	0	0	0	D1
	0	0	0	0	0.5	0	0	0.035	1.0	0	0	D2
	0	0	0	0	0	0.95	0	0	0	0.985	0	D3
	0	0	0	0	0	0	0.8	0	0	0.01	0.999	D4

Original population [P_o] $TP_o = P_1$ $TP_1 = TTP_o = P_2$ $TP_2 = TTTP_o = P_3$ $T^{10}P_o = P_{10}$

$$
\begin{bmatrix} 100 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\quad
\begin{bmatrix} 0 \\ 60 \\ 40 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\quad
\begin{bmatrix} 0 \\ 6 \\ 2 \\ 42 \\ 12 \\ 18 \\ 20 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\quad
\begin{bmatrix} 0 \\ 0.6 \\ 0.1 \\ 6.3 \\ 7.2 \\ 1.8 \\ 5 \\ 39.9 \\ 6 \\ 17.1 \\ 16 \end{bmatrix}
\quad
\begin{bmatrix} 0 \\ 0.207 \\ 0.000 \\ 0.549 \\ 0.145 \\ 0.094 \\ 0.028 \\ 35.734 \\ 23.365 \\ 17.591 \\ 22.288 \end{bmatrix}
$$

Population after n cycles = $T^n P_o$

Consideration of population doubling each time interval:

$$P_1 = T(2P_o)$$
$$P_2 = T(2P_1)$$
$$= T(2(T(2P_o)))$$
$$= 2^2T^2P_o$$
$$= (2T)^2P_o$$
$$P_n = (2T)^nP_o$$

*Something I learned while setting up the introductory examples: There are right and left stochastic matrices. In a right stochastic matrix, the rows add up to 1 and in a left stochastic matrix, the columns add up to 1 [10]. After considering this bit of information and how the product of the stochastic matrix and the initial state matrix yields the outcome state, I figure out the following: If your stochastic n x n matrix, A, is a *left* matrix, you multiply it on the *left* of the matrix of the initial state, B, a n x 1 matrix, to get the outcome state AB. If your stochastic n x n matrix, A, is a *right* matrix, you multiply it on the *right* of the matrix of the initial state, B, a 1 x n matrix, to get the outcome state BA. I initially set up a right matrix and multiplied it on the left of the initial 2 x 1 initial state matrix. When the sum of the entries of the outcome matrix wasn't the same as the sum of the entries of the initial state matrix, I knew something was wrong.

Works Cited

[1] Wikipedia, Andrey Markov, https://en.wikipedia.org/wiki/Andrey_Markov (updated October 8, 2017).

[2] WolframAlpha, Stochastic process, definition used as a word, http://www.wolframalpha.com/input/?i=stochastic+process&rawformassumption=%7B%22C%22,+%22stochastic+process%22%7D+-%3E+%7B%22Word%22%7D (accessed November 12, 2017).

[3] WolframAlpha, Markov process, definition used as a word, http://www.wolframalpha.com/input/?i=Markov+process&lk=1&rawformassumption=%22ClashPrefs%22+-%3E+%7B%22Word%22,+%22Markov+process%22%7D (accessed November 12, 2017).

[4] B. Husen, Markov Processes, https://people.math.osu.edu/husen.1/teaching/571/markov_1.pdf (accessed November 12, 2017).

[5] L.J.S Allen, Tutorial: Stochastic Modeling in Biology - Applications of Discrete- Time Markov Chains, http://www.nimbios.org/tutorials/TT_stochastic_modeling_talks/NIMBioSTut_LJSA_1.pdf (accessed November 12, 2017).

[6] OpenStax College, *Anatomy and Physiology*, Rice University, 2013.

[7] A.C. de Luca, Adult Stem Cells, https://www.fastbleep.com/biology-notes/32/852 (accessed November 14, 2017)

[8] P.A. Tsonis, Stem Cells from Differentiated cells, *Molecular Intervention,* April 2004 4:81-83.

[9] A.K. Chan, Esquema da especificação cellular, https://commons.wikimedia.org/wiki/File:Esquema_da_especifica%C3%A7%C3%A3o_celular.png (accessed November 14, 2017).

[10] Wikipedia, Stochastic matrix, https://en.wikipedia.org/wiki/Stochastic_matrix (updated October 19, 2017).

Using Inverse Functions as an Aid to Compute a Derivative Value

Nick Goins

For the first edition of Continuum, I wrote an article illustrating how to use conjugates to compute the derivative of a radical function, for any index. In what follows, I will present an alternative, algebraic, procedure to compute these same derivatives. This technique will also apply to a wider class of functions, but will prove most useful in the case of radicals. As we will see, the method will require that the point of interest be on the x-axis. However, we will then generalize the method by showing that we can shift a function vertically so that the corresponding point is on the x-axis, before applying the method.

Consider the limit of a difference quotient definition of the derivative,

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

When working with this definition, one of the potential challenges is dealing with the $f(a + h)$ term. In what follows, I will discuss a technique that I discovered in which we will use inverse functions as an aid to compute a derivative value, at a specific point. First, consider the following theorem.

Theorem 1: For a differentiable function $k(x)$, in which $k(0) = 0$, and for a function $f(x)$ which is differentiable at $x = a$, then

$$f'(a) = \lim_{h \to 0} \frac{f(a + k(h)) - f(a)}{k(h)}$$

Proof: Using L'Hopital's Rule,

$$\lim_{h \to 0} \frac{f(a + k(h)) - f(a)}{k(h)} = \lim_{h \to 0} \frac{f'(a + k(h))k'(h) - 0}{k'(h)} = \lim_{h \to 0} f'(a + k(h))$$

Then, interchanging the limit and the function f' and using the fact that $k(h)$ is continuous and passes through the origin,

$$= f'\left(\lim_{h \to 0} a + k(h)\right) = f'(a + k(0)) = f'(a)$$

$\square$

The next example will illustrate the theorem.

Example 1: Let $f(x) = x^2 + 4x$ and let $k(x) = \sin x$. Compute $f'(2)$.

Solution: Using the formula in the above theorem, we get

$$f'(2) = \lim_{h \to 0} \frac{f(2 + k(h)) - f(2)}{k(h)} = \lim_{h \to 0} \frac{f(2 + \sin h) - f(2)}{\sin h} = \lim_{h \to 0} \frac{(2 + \sin h)^2 + 4(2 + \sin h) - 12}{\sin h}$$

$$= \lim_{h \to 0} \frac{4 + 4\sin h + \sin^2 h + 8 + 4\sin h - 12}{\sin h} = \lim_{h \to 0} \frac{8\sin h + \sin^2 h}{\sin h} = \lim_{h \to 0} 8 + \sin h = 8$$

□

Regarding the formula in the theorem, if $a + k(h)$ happened to be the inverse of $f(h)$, then the first term in the difference quotient would reduce to h and so we would get

$$f'(a) = \lim_{h \to 0} \frac{f(a + k(h)) - f(a)}{k(h)} = \lim_{h \to 0} \frac{f(f^{-1}(h)) - f(a)}{k(h)} = \lim_{h \to 0} \frac{h - f(a)}{k(h)}$$

In order for this limit to exist, we would need $f(a)$ to be 0. That is, $x = a$ must be an x-intercept of the function $f(x)$.

Example 2: Compute $f'(0)$, where $f(x) = \sin^{-1} x$.

Solution: Using the formula in the theorem, the value of the derivative is

$$f'(0) = \lim_{h \to 0} \frac{f(0 + k(h)) - f(0)}{k(h)} = \lim_{h \to 0} \frac{f(k(h))}{k(h)} = \lim_{h \to 0} \frac{\sin^{-1}(k(h))}{k(h)}$$

If we choose $k(h) = \sin h$, then $k(h)$ is the inverse function of f, is differentiable at 0 and $k(0) = 0$. Then,

$$= \lim_{h \to 0} \frac{\sin^{-1}(\sin h)}{\sin h} = \lim_{h \to 0} \frac{h}{\sin h} = 1$$

□

In the remaining examples, we will set up the limit of a difference quotient and then find the necessary expression for $k(h)$.

Example 3: Compute $f'(0)$, where $f(x) = \sqrt[3]{x}$.

Solution:

$$f'(0) = \lim_{h \to 0} \frac{f(0 + k(h)) - f(0)}{k(h)} = \lim_{h \to 0} \frac{f(k(h))}{k(h)} = \lim_{h \to 0} \frac{\sqrt[3]{k(h)}}{k(h)}$$

With the goal of having the numerator reduce to h, we will choose $k(h) = h^3$. Also, for this choice of $k(h)$, we see that it has the property of being differentiable and passes through the origin. Then,

$$\lim_{h \to 0} \frac{\sqrt[3]{k(h)}}{k(h)} = \lim_{h \to 0} \frac{\sqrt[3]{h^3}}{h^3} = \lim_{h \to 0} \frac{h}{h^3} = \lim_{h \to 0} \frac{1}{h^2} = \infty$$

□

A Slight Generalization of the Above Technique

If $f(x)$ is differentiable and invertible but $x = a$ is not an x-intercept of $f(x)$ then define $g(x) = f(x) - f(a)$. Then, $g'(a) = f'(a)$ and $x = a$ is an x-intercept of $g(x)$. In other words, if we have every requirement for the above technique except for the fact that $x = a$ is not an x-intercept of $f(x)$, then we can shift the function vertically so that we do get an x-intercept since vertical shifts do not alter the derivative values at the corresponding points. This follows directly from the fact that the derivative of a constant is 0, so that if $g(x) = f(x) + c$, then $g'(a) = f'(a)$.

Example 4: Compute $f'(1)$, where $f(x) = \sqrt[3]{x}$.

Solution: Since $x = 1$ is not an x-intercept of $f(x)$, we will shift the graph down one unit and find the derivative at $x = 1$ of the shifted function. That is, let $g(x) = f(x) - 1$, so that $g(1) = 0$ and also $g'(1) = f'(1)$.

$$f'(1) = g'(1) = \lim_{h \to 0} \frac{g\big(1 + k(h)\big) - g(1)}{k(h)} = \lim_{h \to 0} \frac{g(1 + k(h))}{k(h)}$$

We want to choose $k(h)$ so that $1 + k(h)$ is the inverse of g. Equivalently, we want the numerator, $g(1 + k(h))$, to reduce to h. That is, we want $\sqrt[3]{1 + k(h)} - 1 = h$. Solving this equation for $k(h)$, gives $k(h) = (h + 1)^3 - 1$. Notice that $k(h)$ is differentiable at $h = 0$, and $k(0) = 0$. Then,

$$\lim_{h \to 0} \frac{g\big(1 + k(h)\big)}{k(h)} = \lim_{h \to 0} \frac{h}{(h + 1)^3 - 1} = \lim_{h \to 0} \frac{h}{h^3 + 3h^2 + 3h + 1 - 1}$$

$$= \lim_{h \to 0} \frac{h}{h(h^2 + 3h + 3)} = \lim_{h \to 0} \frac{1}{h^2 + 3h + 3} = \frac{1}{3}$$

That is, $f'(1) = \frac{1}{3}$, as expected.

□

In example 4 if the index was larger than 3, say n, then expanding the term $(h + 1)^n$ would've been more work. However, notice that when we evaluated the limit we really only needed the second to last term (why?). The following will prove to be very useful in the case of radicals with a larger index.

Note: In the remaining examples, we are going to use the binomial theorem, in the following way,

$$(h + a)^n = hp(h) + a^n$$

That is, expanding $(h + a)^n$ yields a sum of terms in which all but the last have a factor of h. We can factor out the h from the first terms, leaving a polynomial $p(h)$, and the last term is a^n. The details for this are as follows,

$$(h + a)^n = \sum_{k=0}^{n} \binom{n}{k} h^k a^{n-k} = \sum_{k=1}^{n} \binom{n}{k} h^k a^{n-k} + a^n = h \sum_{k=1}^{n} \binom{n}{k} h^{k-1} a^{n-k} + a^n$$

Then, define $p(h) = \sum_{k=1}^{n} \binom{n}{k} h^{k-1} a^{n-k}$. In particular, we can write $p(h)$ as follows,

$$p(h) = \sum_{k=1}^{n} \binom{n}{k} h^{k-1} a^{n-k} = \binom{n}{1} h^{1-1} a^{n-1} + \sum_{k=2}^{n} \binom{n}{k} h^{k-1} a^{n-k} = na^{n-1} + \sum_{k=2}^{n} \binom{n}{k} h^{k-1} a^{n-k}$$

so that

$$\boxed{p(0) = na^{n-1}}$$

The function value $p(0)$ is the coefficient of the second to last term in the polynomial $p(h)$. In example 4, we could've written $(h + 1)^3 = hp(h) + 1^3 = hp(h) + 1$, and then $p(0) = 3(1)^2 = 3$. This will become more useful in the following examples, as we consider expansions with powers larger than 3.

Example 5: Compute $f'(3)$, where $f(x) = \sqrt[7]{x}$.

Solution: Since $x = 3$ does not correspond to an x-intercept of $f(x)$, we define $g(x) = f(x) - f(3)$. That is, $g(x) = \sqrt[7]{x} - \sqrt[7]{3}$. Then,

$$f'(3) = g'(3) = \lim_{h \to 0} \frac{g(3 + k(h)) - g(3)}{k(h)} = \lim_{h \to 0} \frac{g(3 + k(h))}{k(h)} = \lim_{h \to 0} \frac{\sqrt[7]{3 + k(h)} - \sqrt[7]{3}}{k(h)}$$

The goal in finding $k(h)$ is that we want the numerator to reduce to h. Again, this is equivalent to $3 + k(h) = g^{-1}(h)$, so we solve $\sqrt[7]{3 + k(h)} - \sqrt[7]{3} = h$, for $k(h)$. Doing so, gives

$$k(h) = \left(h + \sqrt[7]{3}\right)^7 - 3$$

Substituting this into the above limit gives

$$\lim_{h \to 0} \frac{\sqrt[7]{3 + k(h)} - \sqrt[7]{3}}{k(h)} = \lim_{h \to 0} \frac{\sqrt[7]{3 + \left(h + \sqrt[7]{3}\right)^7 - 3} - \sqrt[7]{3}}{\left(h + \sqrt[7]{3}\right)^7 - 3} = \lim_{h \to 0} \frac{h}{\left(h + \sqrt[7]{3}\right)^7 - 3}$$

To evaluate this limit using algebra, we could expand the seventh power, but that is much more work than we need to do. We will use the observation, $(h + a)^n = hp(h) + a^n$. That is,

$$\lim_{h \to 0} \frac{h}{\left(h + \sqrt[7]{3}\right)^7 - 3} = \lim_{h \to 0} \frac{h}{hp(h) + \left(\sqrt[7]{3}\right)^7 - 3}$$

Reducing this expression,

$$= \lim_{h \to 0} \frac{h}{hp(h) + 3 - 3} = \lim_{h \to 0} \frac{h}{hp(h)} = \lim_{h \to 0} \frac{1}{p(h)} = \frac{1}{p(0)}$$

In the discussion above, we found the function value $p(0) = na^{n-1}$, where n is the index and a is the value at which we are finding the derivative, so that we get the following final result

$$f'(3) = \frac{1}{7\left(\sqrt[7]{3}\right)^6}$$

□

Note: In example 3, $f(k(h)) = h$, so that $k(h) = f^{-1}(h)$. In example 4, $x = 1$ was not an x-intercept of $f(x)$, so we shifted the function vertically downwards one unit. Then in that example, we found $g(1 + k(h)) = h$, so that $1 + k(h) = g^{-1}(h)$ and so $k(h) = g^{-1}(h) - 1$. In general, the relationship between the original function $f(x)$, the value of a and the function $k(h)$, is expressed as follows

$$\boxed{k(h) = f^{-1}(h + f(a)) - a}$$

In example 5, we found $k(h) = \left(h + \sqrt[7]{3}\right)^7 - 3$, which has the same form as the boxed formula stated above. If $x = a$ corresponds to an x-intercept of $f(x)$ at the origin (so that $a = 0$ and $f(a) = 0$), then $k(h)$ is exactly the inverse of f. If $x = a$ is an x-intercept of $f(x)$ but not at the origin, then $k(h)$ is a vertical shift of the inverse of f. If $a = 0$, but $f(a) \neq 0$, then $k(h)$ is a horizontal shift of f^{-1}. Of course, a need not be zero and need not correspond to an x-intercept. The same expression works for $k(h)$ in each scenario. In what follows, we won't use the above expression for $k(h)$, but will find it as we progress through the examples. The overall goal is to illustrate the algebra that is needed to evaluate these derivative values, not to find simplifying formulas.

Example 6: Compute $f'(11)$, where $f(x) = \sqrt[19]{x}$.

Solution: Since $x = 11$ does not correspond to an x-intercept of $f(x)$, we define $g(x) = f(x) - f(11)$. That is, $g(x) = \sqrt[19]{x} - \sqrt[19]{11}$. Then,

$$f'(11) = g'(11) = \lim_{h \to 0} \frac{g(11 + k(h)) - g(11)}{k(h)} = \lim_{h \to 0} \frac{g(11 + k(h))}{k(h)} = \lim_{h \to 0} \frac{\sqrt[19]{11 + k(h)} - \sqrt[19]{11}}{k(h)}$$

Similar to the previous example, $k(h) = \left(h + \sqrt[19]{11}\right)^{19} - 11$, so that

$$\lim_{h\to 0}\frac{\sqrt[19]{11+k(h)}-\sqrt[19]{11}}{k(h)}=\lim_{h\to 0}\frac{\sqrt[19]{11+\left(h+\sqrt[19]{11}\right)^{19}-11}-\sqrt[19]{11}}{\left(h+\sqrt[19]{11}\right)^{19}-11}=\lim_{h\to 0}\frac{h}{\left(h+\sqrt[19]{11}\right)^{19}-11}$$

$$=\lim_{h\to 0}\frac{h}{hp(h)+\left(\sqrt[19]{11}\right)^{19}-11}=\lim_{h\to 0}\frac{h}{hp(h)+11-11}$$

$$=\lim_{h\to 0}\frac{h}{hp(h)}=\lim_{h\to 0}\frac{1}{p(h)}=\frac{1}{p(0)}=\frac{1}{19\left(\sqrt[19]{11}\right)^{18}}$$

□

The method presented above works for functions other than radicals, as illustrated by the next example.

Example 7: Compute $f'(5)$, where $f(x)=\ln x$.

Solution: First, define $g(x)=f(x)-f(5)$.

$$g'(5)=\lim_{h\to 0}\frac{g(5+k(h))-g(5)}{k(h)}=\lim_{h\to 0}\frac{g(5+k(h))}{k(h)}=\lim_{h\to 0}\frac{\ln(5+k(h))-\ln 5}{k(h)}$$

Solving the equation $\ln(5+k(h))-\ln 5=h$, gives

$$k(h)=e^{h+\ln 5}-5$$

then substitution back into the limit,

$$\lim_{h\to 0}\frac{\ln(5+k(h))-\ln 5}{k(h)}=\lim_{h\to 0}\frac{\ln\left(5+e^{h+\ln 5}-5\right)-\ln 5}{e^{h+\ln 5}-5}=\lim_{h\to 0}\frac{h}{e^{h+\ln 5}-5}$$

The expression inside the limit can be manipulated as follows

$$=\lim_{h\to 0}\frac{h}{e^h e^{\ln 5}-5}=\lim_{h\to 0}\frac{h}{5e^h-5}=\lim_{h\to 0}\frac{h}{5(e^h-1)}=\frac{1}{5}\lim_{h\to 0}\frac{1}{\left(\frac{e^h-1}{h}\right)}=\frac{1}{5}\frac{1}{\lim_{h\to 0}\left(\frac{e^h-1}{h}\right)}=\frac{1}{5}$$

□

The preceding examples can be summarized by the following theorem, which is only stated for completeness, and will not be used in any of the remaining examples.

Theorem 2: For a function $f(x)$ which is differentiable at $x=a$ and is invertible,

$$f'(a) = \lim_{h \to 0} \frac{h}{f^{-1}(h + f(a)) - a}$$

Proof: The limit has indeterminate form $\frac{0}{0}$, so we can apply L'Hopital's Rule,

$$\lim_{h \to 0} \frac{h}{f^{-1}(h + f(a)) - a} = \lim_{h \to 0} \frac{\frac{d}{dh}[h]}{\frac{d}{dh}[f^{-1}(h + f(a)) - a]} = \lim_{h \to 0} \frac{1}{(f^{-1})'(h + f(a))}$$

Then, using the formula for the derivative of an inverse,

$$= \lim_{h \to 0} \frac{1}{\left[\frac{1}{f'(f^{-1}(h + f(a)))}\right]}$$

Since f is differentiable, then f' is continuous and since f is invertible, we know that f^{-1} is continuous as well. Thus, we can interchange the operation of taking a limit with these two functions,

$$= \frac{1}{\left[\frac{1}{f'\left(f^{-1}\left(\lim_{h \to 0} h + f(a)\right)\right)}\right]} = \frac{1}{\left[\frac{1}{f'(f^{-1}(f(a)))}\right]} = \frac{1}{\left[\frac{1}{f'(a)}\right]} = f'(a)$$

□

In the remaining examples, we will consider functions built by composition. That is, the "inside" function is not just x.

Example 8: Compute $f'(3)$, where $f(x) = \sqrt[5]{2x + 1}$.

Solution: Since $x = 3$ is not an x-intercept of $f(x)$, we define $g(x) = f(x) - f(3)$. Then, the derivative of $f(x)$ evaluated at $x = 3$ is

$$f'(3) = g'(3) = \lim_{h \to 0} \frac{g(3 + k(h)) - g(3)}{k(h)} = \lim_{h \to 0} \frac{g(3 + k(h))}{k(h)} = \lim_{h \to 0} \frac{\sqrt[5]{2(3 + k(h)) + 1} - \sqrt[5]{7}}{k(h)}$$

To find the expression for $k(h)$, set the numerator equal to h and solve for $k(h)$.

$$\sqrt[5]{2(3 + k(h)) + 1} - \sqrt[5]{7} = h \quad \Rightarrow \quad 2k(h) + 7 = \left(h + \sqrt[5]{7}\right)^5 \quad \Rightarrow \quad k(h)$$
$$= \frac{1}{2}\left(h + \sqrt[5]{7}\right)^5 - \frac{7}{2}$$

Substituting this expression into the limit,

$$\lim_{h\to0}\frac{\sqrt[5]{2\left[\frac{1}{2}\left(h+\sqrt[5]{7}\right)^5-\frac{7}{2}\right]+7}-\sqrt[5]{7}}{\frac{1}{2}\left(h+\sqrt[5]{7}\right)^5-\frac{7}{2}}=\lim_{h\to0}\frac{\sqrt[5]{\left(h+\sqrt[5]{7}\right)^5-7+7}-\sqrt[5]{7}}{\frac{1}{2}\left[\left(h+\sqrt[5]{7}\right)^5-7\right]}=2\lim_{h\to0}\frac{\sqrt[5]{\left(h+\sqrt[5]{7}\right)^5}-\sqrt[5]{7}}{\left(h+\sqrt[5]{7}\right)^5-7}$$

$$=2\lim_{h\to0}\frac{h}{hp(h)+\left(\sqrt[5]{7}\right)^5-7}=2\lim_{h\to0}\frac{1}{p(h)}=\frac{2}{p(0)}=\frac{2}{5\left(\sqrt[5]{7}\right)^4}$$

$\square$

Note: Consider the expression, $((h+a)^m+b)^n$, for positive integers m and n. Using a previously discussed formula, we can write this as follows

$$((h+a)^m+b)^n=(hp(h)+a^m+b)^n$$

Using the binomial theorem, this can then be written as follows

$$(hp(h)+a^m+b)^n=\left(hp(h)+(a^m+b)\right)^n=\sum_{i=0}^n\binom{n}{i}(hp(h))^i(a^m+b)^{n-i}$$

Then pull off the $i=0$ term and factor an h out of the sum,

$$=\left[\sum_{i=1}^n\binom{n}{i}h^ip^i(h)(a^m+b)^{n-i}\right]+(a^m+b)^n=h\left[\sum_{i=1}^n\binom{n}{i}h^{i-1}p^i(h)(a^m+b)^{n-i}\right]+(a^m+b)^n$$

Label the summation in brackets in the latter expression as $q(h)$. The following will show how to compute the value of $q(0)$.

$$q(h)=\sum_{i=1}^n\binom{n}{i}h^{i-1}p^i(h)(a^m+b)^{n-i}$$

$$=\binom{n}{1}h^{1-1}p^1(h)(a^m+b)^{n-1}+\sum_{i=2}^n\binom{n}{i}h^{i-1}p^i(h)(a^m+b)^{n-i}$$

$$=np(h)(a^m+b)^{n-1}+h\sum_{i=2}^n\binom{n}{i}h^{i-2}p^i(h)(a^m+b)^{n-i}$$

Thus,

$$q(0) = np(0)(a^m + b)^{n-1}$$

where

$$((h + a)^m + b)^n = hq(h) + (a^m + b)^n$$

Example 9: Compute $f'(1)$, where $f(x) = \sqrt{\sqrt[3]{x} + 4}$.

Solution: Define $g(x) = f(x) - f(1)$. Then,

$$f'(1) = g'(1) = \lim_{h\to 0} \frac{g(1 + k(h)) - g(1)}{k(h)} = \lim_{h\to 0} \frac{g(1 + k(h))}{k(h)} = \lim_{h\to 0} \frac{\sqrt{\sqrt[3]{1 + k(h)} + 4} - \sqrt{5}}{k(h)}$$

Using the same idea as the previous examples, we solve $\sqrt{\sqrt[3]{1 + k(h)} + 4} - \sqrt{5} = h$, for $k(h)$. Then,

$$k(h) = \left(\left(h + \sqrt{5}\right)^2 - 4\right)^3 - 1$$

Which gives the following

$$\lim_{h\to 0} \frac{\sqrt{\sqrt[3]{1 + k(h)} + 4} - \sqrt{5}}{k(h)} = \lim_{h\to 0} \frac{h}{\left(\left(h + \sqrt{5}\right)^2 - 4\right)^3 - 1}$$

Notice that the denominator has the same form as the expressions discussed in the note preceding this example. Taking advantage of that

$$\lim_{h\to 0} \frac{h}{\left(\left(h + \sqrt{5}\right)^2 - 4\right)^3 - 1} = \lim_{h\to 0} \frac{h}{hq(h) + \left(\left(\sqrt{5}\right)^2 + (-4)\right)^3 - 1} = \lim_{h\to 0} \frac{h}{hq(h)} = \frac{1}{q(0)}$$

Finally, using the previously discovered value of $q(0)$, we get the final result

$$f'(1) = \frac{1}{3 \cdot 2\left(\sqrt{5}\right)^{2-1} \left(\left(\sqrt{5}\right)^2 + (-4)\right)^{3-1}} = \frac{1}{6\sqrt{5}}$$

□

Example 10: Compute $f'(4)$, where $f(x) = \sqrt[3]{x^2 + 1}$.

Solution: Define $g(x) = f(x) - f(4)$. Then,

$$f'(4) = g'(4) = \lim_{h \to 0} \frac{g(4 + k(h)) - g(4)}{k(h)} = \lim_{h \to 0} \frac{g(4 + k(h))}{k(h)} = \lim_{h \to 0} \frac{\sqrt[3]{(4 + k(h))^2 + 1} - \sqrt[3]{17}}{k(h)}$$

As the reader can verify, we will choose $k(h) = \sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} - 4$, so that the above limit becomes

$$= \lim_{h \to 0} \frac{h}{\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} - 4}$$

To evaluate this limit, we will use a conjugate,

$$= \lim_{h \to 0} \frac{h}{\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} - 4} \cdot \frac{\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4}{\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4} = \lim_{h \to 0} \frac{h\left(\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4\right)}{\left(h + \sqrt[3]{17}\right)^3 - 1 - 16}$$

Using the binomial theorem again,

$$= \lim_{h \to 0} \frac{h\left(\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4\right)}{hp(h) + \left(\sqrt[3]{17}\right)^3 - 17} = \lim_{h \to 0} \frac{h\left(\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4\right)}{hp(h)} = \lim_{h \to 0} \frac{\sqrt{\left(h + \sqrt[3]{17}\right)^3 - 1} + 4}{p(h)} = \frac{8}{p(0)}$$

In the above example, $p(0) = na^{n-1} = 3\left(\sqrt[3]{17}\right)^2$, so that $f'\left(\sqrt[3]{17}\right) = \frac{8}{3\left(\sqrt[3]{17}\right)^2}$.

$\square$

In example 10, we used the conjugate for a square root function because of the squared term in the radicand. If that exponent was a larger integer, say n, then we would end up having to use a conjugate for a radical with index n. This method is presented in the first edition of Continuum and the final example will illustrate using the method of this article along with the method from that aforementioned article.

Example 11: Compute $f'(2)$, where $f(x) = \sqrt[7]{x^4 - 5}$.

Solution: Define $g(x) = f(x) - f(2)$. Then,

$$f'(2) = g'(2) = \lim_{h \to 0} \frac{g(2 + k(h)) - g(2)}{k(h)} = \lim_{h \to 0} \frac{g(2 + k(h))}{k(h)}$$

$$= \lim_{h \to 0} \frac{\sqrt[7]{(2 + k(h))^4 - 5} - \sqrt[7]{11}}{k(h)}$$

Solving for $k(h)$, gives $k(h) = \sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} - 2$. Then, substituting this back into the above limit gives

$$= \lim_{h \to 0} \frac{\sqrt[7]{\left(2 + \sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} - 2\right)^4 - 5} - \sqrt[7]{11}}{\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} - 2} = \lim_{h \to 0} \frac{h}{\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} - 2}$$

To evaluate this limit, we need to use a conjugate for a radical of index 4 (see the first edition of Continuum). To find the appropriate conjugate, consider the following factorization

$$a^4 - b^4 = (a - b)(a + b)(a^2 + b^2)$$

So, if the denominator in the above limit is "$a - b$" and if we want to take that fourth root to the fourth power, we will multiply by $(a + b)(a^2 + b^2)$. That is,

$$= \lim_{h \to 0} \frac{h}{\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} - 2} \cdot \frac{\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} + 2\right)\left(\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5}\right)^2 + 2^2\right)}{\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} + 2\right)\left(\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5}\right)^2 + 2^2\right)}$$

$$= \lim_{h \to 0} \frac{h\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} + 2\right)\left(\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5}\right)^2 + 2^2\right)}{\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5}\right)^4 - 2^4}$$

$$= \lim_{h \to 0} \frac{h\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5} + 2\right)\left(\left(\sqrt[4]{\left(h + \sqrt[7]{11}\right)^7 + 5}\right)^2 + 2^2\right)}{\left(h + \sqrt[7]{11}\right)^7 + 5 - 16}$$

$$= \lim_{h \to 0} \frac{h\left(\sqrt[4]{hp(h) + (\sqrt[7]{11})^7 + 5} + 2\right)\left(\left(\sqrt[4]{hp(h) + (\sqrt[7]{11})^7 + 5}\right)^2 + 2^2\right)}{hp(h) + (\sqrt[7]{11})^7 - 11}$$

$$= \lim_{h \to 0} \frac{\left(\sqrt[4]{hp(h) + 16} + 2\right)\left(\left(\sqrt[4]{hp(h) + 16}\right)^2 + 4\right)}{p(h)}$$

$$= \frac{\left(\sqrt[4]{16} + 2\right)\left(\left(\sqrt[4]{16}\right)^2 + 4\right)}{p(0)}$$

$$= \frac{4 \cdot 8}{7(\sqrt[7]{11})^6}$$

$$= \frac{32}{7(\sqrt[7]{11})^6}$$

Johann Bernoulli

Born in Basel, Swiss Confederacy, on July 27th, 1667, seventeen years before Leibniz would publish his earth-shattering paper on calculus, Johann was the younger brother and lifelong rival of his equally famed mathematician brother, Jakob. (Jakob introduced the term integral in 1690; Johann claimed it was his coinage.) His father wanted him to be a merchant, but he was excited by mathematics. His first work in mathematics was *Solutio problematis funicularii*, which appeared in *Acta eruditorum* (the same journal Leibniz's paper appeared in) in 1690.

Johann made fun of his brother frequently. When Jakob though he had correctly described the shape of a sail, pointing out that in different circumstances it might take different shapes, Johann seems to think he is claiming that the curve which describes the shape of one sail is made up of parts of other curves [Peiffer]. When Jakob shared with Johann the differential equation that yields the answer, Johann solved it and said: "Once more he forces me to complete the solution that he has begun and developed until this equation, after which he apparently gave up" (*Journal des savants*, 1692, p. 189).

It seems likely that Johann did not want to be outshone by his brother, or always considered as secondary to the great Jakob. In a letter to Leibniz, Jakob claims to have been the first to understand Leibniz's calculus. When Jakob found a formula for curvature, Johann printed it on cards and passed them out as his business card! Jakob then translated his formula into polar coordinates and announced it as "*inconnu même de mon frère*" – "unknown even by my brother." [Peiffer]

Johann wrote to L'Hopital on January 12, 1695, speaking of his brother: "He is a misanthropist in general and does not even spare his own brother...He is filled with rage, hate, envy and jealousy against me." In a later letter: "You would not believe how much this brother, unworthy of the name, hates me, persecutes me and tries to destroy me." [Peiffer]

In Paris through most of 1691, Johann taught calculus to the Marquis de L'Hopital, who wrote the first book on the subject in 1696, *Analyse des infiniments petits*.

Johann became a Doctor of Medicine in 1694 when he applied differential calculus to muscular contractions.

Johann proposed a problem, called the brachistochrone problem, in *Acta eruditorum* in June 1696: what is the curve of quickest descent of a body moving from a higher point to a lower point in a plane; or, from *Wolfram Mathworld*, "Find the shape of the curve down which a bead sliding from rest and accelerated by gravity will slip (without friction) from one point to another in the least time. The term derives from the Greek βραχιστοσ (*brachistos*) "the shortest" and χρονοσ (*chronos*) "time, delay." "

Five solutions were published in the magazine almost a whole year later, in May, 1697. One of them appeared anonymously.

At this time, Newton was retired, and he was working in the Royal Mint. He got home one day and saw the problem in a letter from Johann. It had been a long day. He was tired. But he sat down and solved the problem. Right then and there, in one night. He sent his answer to the magazine anonymously. It had taken Bernoulli two weeks to solve the problem

When Bernoulli read the anonymous solution, he knew right away that Isaac Newton had come up with it. How did he know? Well, what he said was,

"I know the lion by his paw."

Johann arrived at a solution by thinking about a beam of light. He knew, according to Fermat's Principle, that whatever path between two points requires the least time, is the path that a beam of light will always follow. He imagined how a beam of light would move if affected by gravity.

Leibniz took one week to solve the problem.

In 1700, Johann's son Daniel was born. Daniel would also become a mathematician, and a friend of the great Euler. According to E.T. Bell [Men of Mathematics] Daniel was "kicked out of the house" for winning a

prize which his father had also competed for. Johann also stole one of his son's papers, changed the name and date, and submitted it as his own work. Nikolaus and Johann II were the other sons of Johann Bernoulli, and they too were fine mathematicians.

One of the most important activities of Johann was his fierce partisanship of Leibniz in the calculus priority dispute. "Johann saw himself as Horatio, bravely defending Leibnizian calculus from the arrogant, misguided English." [Tent] Johann fostered the idea that Newton had plagiarized Leibniz, and in a letter of 1713 he attacked Newton's character – he later denied having written the letter. Johann encouraged Leibniz to publish "challenge problems" which he felt the Newtonian mathematicians would be unable to solve with what he believed were their inferior methods. "Newton would, as you know, find himself in difficulties." (Hall, p. 216). The French Academy of Sciences took Leibniz's side, printing an attack by Bernoulli on Newton's *Principia* without allowing a response. (Hall, p. 214). After Leibniz's death, Bernoulli did try to mend fences with Newton, writing "I desire nothing so much as to live in good fellowship with him, and to find an opportunity of showing him how much I value his rare merits, indeed I never speak of him save with much praise." (Hall, p. 238).

One of Johann's greatest contributions was what he referred to as "exponential calculus" (now considered just a part of calculus) which contained the rule $d(\ln u) = du/u$. However, Johann also believed that logs of negative numbers existed and were real numbers. "He drew this remarkable, counterintuitive conclusion by applying what we would call the chain rule in taking the derivative of $\ln(-x)$." [Bedard]

Another tremendously important contribution of Johann Bernoulli to mathematics is an indirect one: he was the tutor of the great Leonhard Euler. Euler honored Bernoulli, calling him "illustrious."

Bedard, Paul: "Euler and the Bernoullis: Learning by Teaching" *MAA Convergence* August 2015.

Bell, E.T: Men of Mathematics, Simon & Schuster, 1937

Hall, Alfred Rupert: Philosophers at War: The Quarrel Between Newton and Leibniz, Cambridge University Press, 1980.

Peiffer,Jean: "Jacob Bernoulli, Teacher and rival of his brother Johann," *Journal Electronique d'Histoire de Probabilite et de la Statistique*, Vol 2 No. 1 November 2006)

Tent, M.B.W. : Leonhard Euler and the Bernoullis: Mathematicians from Basel, Taylor and Francis Group, 2009.

Questions

1.   List 5 noted mathematicians from the Bernoulli family.

2.   What seems to have been the attitude of Bernoulli towards the intellectual property of others?

3.   What is the name (no longer used) of the branch of calculus which Johann Bernoulli was a pioneer in?

4.   What error did Johann make about logarithms?

5.   Applying the chain rule and ignoring the domain issues, find the (nonexistent) derivative of $y = \ln(-x),\ x > 0$

6.   How did Johann sole the brachistochrone problem?

7.   Johann said of Newton, "I never speak of him save with much praise." What is one quote which belies this claim?

8.   How did Johann hope to humiliate British mathematicians?

9.   Other than Euler, name a famous student of Johann Bernoulli.

10.   Who introduced the term *integral*?

11.   What did Johann print on his business cards?

12.   How long did it take Leibniz to solve the brachistochrone problem?

13.   How long did it take Newton to solve the brachistochrone problem?

14.   What earned Johann Bernoulli a doctorate in medicine?

Reparametrization by Arc Length

Nick Goins

For a vector function $\vec{r}(t)$, the input variable t is referred to as a parameter. The word parameter is also used when referring to the arc length parameter, the context should make it clear which of the two uses is intended. In what follows we will discuss the topic of changing the input parameter for a vector function and what impact this has on the graph of the vector function, the description of the points on the graph and of the speed function.

Recall that the arc length, $\int_a^b |\vec{v}(t)|\, dt$ is the length of the curve representing $\vec{r}(t)$ from the point corresponding to $t = a$ to the point corresponding to $t = b$, more precisely, the length of the curve from the terminal point of the vector $\vec{r}(a)$ to the terminal point of the vector $\vec{r}(b)$. The notation, $P(a)$, is frequently used for the terminal point of $\vec{r}(a)$. We can then create a function which measures how far we've travelled from the terminal point of $\vec{r}(a)$ after a time t, by replacing the upper limit of integration with the variable t. We call this function the arc length parameter, $s(t) = \int_a^t |\vec{v}(\tau)|\, d\tau$ and we refer to a as the basepoint. Again, more precisely, the basepoint is the terminal point of the vector $\vec{r}(a)$.

1. Let $\vec{r}(t) = \langle 3t, t + 4 \rangle$. Find the arc length parameter with basepoint corresponding to $t = 0$. Then compute $s(0), s(1)$ and $s(\sqrt{10})$.

2. Let $\vec{r}(t) = \langle t, t + 2 \rangle$. Find the arc length parameter with basepoint corresponding to $t = 2$. Then compute $s(0), s(1)$ and $s(\sqrt{2})$.

Note: For a vector function $\vec{r}$ which is not constant, its arc length parameter satisfies the following

$$ s(t) = \int_{t_0}^t \left| \frac{d\vec{r}(\tau)}{d\tau} \right| d\tau \qquad \Rightarrow \qquad \frac{ds}{dt} = \left| \frac{d\vec{r}(t)}{dt} \right| > 0 $$

That is, the function $s(t)$ is increasing. One implication of this is that $s(t)$ has an inverse function. The following describes the inputs and outputs for $s(t)$ and $s^{-1}(t)$

$$ s: c \longmapsto l $$

$$ s^{-1}: l \longmapsto c $$

That is, the function $s(t)$ gives the arc length on the curve described by $\vec{r}(t)$ from $t = a$ (the basepoint) to the point corresponding to $t = c$. The inverse function takes an arc length l and gives the value of t in which the arc length between $t = a$ and $t = c$ is l. In particular, if we find the arc length parameter with basepoint $P(a)$, then $s(a) = 0$ and so $s^{-1}(0) = a$.

3. Let $\vec{r}(t) = \langle 3t, 2t - 1 \rangle$. Find the arc length parameter with basepoint corresponding to $t = 1$. Then find the inverse of this function. That is, find $s^{-1}(t)$.

4. Let $\vec{r}(t) = \langle 3t, t + 4 \rangle$. Using your work from exercise #1, find the inverse of the arc length parameter. That is, find $s^{-1}(t)$.

5. Let $\vec{r}(t) = \left\langle \frac{2}{9}(3t - 1)^{\frac{3}{2}}, \frac{2}{3}(t - 1)^{\frac{3}{2}} \right\rangle$. Find the arc length parameter with basepoint corresponding to $t = 0$. Then find the inverse of this function. That is, find $s^{-1}(t)$.

Example: Consider the vector functions $\vec{r}(t) = \langle 2t, t + 1 \rangle$ and $\vec{R}(t) = \langle 6t, 3t + 1 \rangle$, which represent lines in the xy-plane. For $t = 1$, the point on the graph for $\vec{r}(t)$ is $(2, 2)$ and the point on the graph for $\vec{R}(t)$ is $(6, 4)$. Similarly, the point on the graph for $\vec{r}(t)$ when $t = 0$ is $(0, 1)$ and the point on the graph for $\vec{R}(t)$ when $t = 0$ is also $(0, 1)$. Thus, the graphs pass through the same point. To find the slope of the line represented by $\vec{r}(t)$, choose two points on the line, such as $(2a, a + 1)$ and $(2b, b + 1)$. The slope through these points is

$$\frac{\Delta y}{\Delta x} = \frac{(b + 1) - (a + 1)}{2b - 2a} = \frac{b - a}{2(b - a)} = \frac{1}{2}$$

Similarly, to find the slope of the line represented by $\vec{R}(t)$, choose two points on the line, such as $(6a, 3a + 1)$ and $(6b, 3b + 1)$. The slope through these points is

$$\frac{\Delta y}{\Delta x} = \frac{(3b + 1) - (3a + 1)}{6b - 6a} = \frac{3(b - a)}{6(b - a)} = \frac{1}{2}$$

Therefore, the lines are the same. Notice the relationship, $\vec{R}(t) = \vec{r}(3t)$.

Definition: For a vector function $\vec{r}(t)$, a reparametrization is another vector function $\vec{R}(t)$ which has the same graph as $\vec{r}(t)$, but the points are labelled differently.

In the above example, the point $(4, 3)$ is on the graph of $\vec{r}(t)$, and corresponds to $t = 2$. The point is also on the graph of $\vec{R}(t)$ and corresponds to $t = \frac{2}{3}$. So, the points are the same, they are just labelled differently depending on which vector function we are referring to.

Constructing a Reparametrization: For a vector function, $\vec{r}(t)$, a reparametrization will have the form $\vec{r}(f(t))$.

Note that there are some restrictions on the function $f(t)$. In particular, the domain of f must be $\mathbb{R}$ and we need the range of f to match the domain of $\vec{r}$. Thus, the domain and range of f must be $\mathbb{R}$. In addition, we want f to be monotonic. For example, if $\vec{r}(t) = \langle t, t^3 \rangle$ and $f(t) = t^2$, then $\vec{r}(f(t)) = \langle t^2, t^6 \rangle$. The graph of $\vec{r}(f(t))$ does not match the graph of $\vec{r}(t)$ (why?).

Example: From the previous example, we can see that $\vec{R}(t) = \langle 6t, 3t + 1 \rangle$ is a reparametrization of $\vec{r}(t) = \langle 2t, t + 1 \rangle$ because $\vec{R}(t) = \vec{r}(f(t))$ where $f(t) = 3t$.

Example: Let $\vec{r}(t) = \langle t^2, e^t \rangle$, $f(t) = 4t$ and $g(t) = -2t + 5$. Then the following are both reparametrizations of $\vec{r}(t)$,

$$\vec{r}(f(t)) = \langle (4t)^2, e^{4t} \rangle = \langle 16t^2, e^{4t} \rangle$$
$$\vec{r}(g(t)) = \langle (-2t + 5)^2, e^{-2t+5} \rangle = \langle 4t^2 - 20t + 25, e^{-2t+5} \rangle$$

Thus, a vector function can have infinitely many different reparametrizations.

6. Let $\vec{r}(t) = \left\langle t, \frac{1}{2}t^2 \right\rangle$.

a) Sketch the graph of $\vec{r}(t)$ by first filling in the following table. Label the point on your graph corresponding to $t = 0$ and $t = 1$.

input	output	point on graph
0		
1		
−1		
2		
−2		

b) Let $f(t) = 3t$. Sketch the graph of $\vec{r}(f(t))$ by first filling in the following table. Label the point on your graph corresponding to $t = 0$ and $t = 1$.

input	output	point on graph
0		
1		
−1		
2		
−2		

7. Let $\vec{r}(t) = \langle t^2 + t, t + 2 \rangle$.

a) Sketch the graph of $\vec{r}(t)$ by first filling in the following table. Label the point on your graph corresponding to $t = 0$ and $t = 1$.

| input	output	point on graph
0		
1		
−1		
2		
−2		
−0.5		

b) Let $f(t) = t + 1$. Sketch the graph of $\vec{r}(f(t))$ by first filling in the following table. Label the point on your graph corresponding to $t = 0$ and $t = 1$.

input	output	point on graph
0		
1		
−1		
2		
−2		
−1.5		

In the above exercises, notice that reparametrizing did not change the graph, it only changed the description of each point, as you illustrated by labelling the points corresponding to $t = 0$ and $t = 1$.

Reparametrization by Arc Length

For a vector function $\vec{r}(t)$ and for a fixed basepoint, $P(a)$ (that is, the point on the graph corresponding to $t = a$), we are going to label a point by its distance along the curve from $P(a)$.

$\vec{r}(t)$

original vector function

$\vec{r}\left(s^{-1}(t)\right)$

vector function reparametrized by arc length
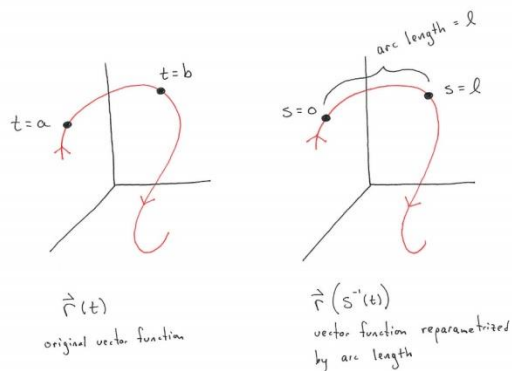
The above provides a visual for the relationship $s(a) = 0$ and so $s^{-1}(0) = a$, that is the same point is either labelled as 0 or as a, depending on the vector function description of the curve.

Definition: For a vector function $\vec{r}(t)$, the speed function is $s'(t) = |\vec{v}|$.

Note: The unit tangent vector is $\vec{T} = \frac{\vec{v}}{|\vec{v}|}$, so that $\vec{v} = |\vec{v}|\vec{T}$. This can then be written as follows

$$\frac{d\vec{r}}{dt} = s'(t)\vec{T}(t)$$

and shows that we can factor the derivative of a vector function into two components. That is, the derivative is the product of the speed and tangent vector which shows that the derivative has a dynamical component and a geometric component.

Procedure to find the Reparametrization by Arc Length:
For the vector function $\vec{r}(t)$ with basepoint corresponding to $t = a$.

1. Find the arc length parameter $s(t)$. That is, $s(t) = \int_a^t |\vec{v}(\tau)|\, d\tau$.
2. Find the inverse function, $s^{-1}(t)$.
3. Simplify the composition, $\vec{R}(t) = \vec{r}\left(s^{-1}(t)\right)$.

8. Let $\vec{r}(t) = \langle 3t, t + 4\rangle$. Reparametrize $\vec{r}(t)$ by arc length, using $t = 0$ as the basepoint for s. Use your work from exercise #4.

9. Let $\vec{r}(t) = \langle 2t + 1, 3t - 1\rangle$. Reparametrize $\vec{r}(t)$ by arc length, using $t = -1$ as the basepoint for s.

10. Let $\vec{r}(t) = \langle 2t, t + 1 \rangle$.

a) Find the point on the graph of $\vec{r}(t)$ corresponding to $t = 1$.

b) Find the arc length parameter $s(t)$, with basepoint corresponding to $t = 0$.

c) Find $s^{-1}(t)$.

d) Compute $s^{-1}(0)$.

e) Compute the arc length of $\vec{r}(t)$ from $t = 0$ to $t = 1$.

f) Reparametrize $\vec{r}(t)$ by arc length. That is, find $\vec{R}(t) = \vec{r}(s^{-1}(t))$.

g) Evaluate $\vec{R}(t)$ using the value of the arc length from part c as the input. That is, the t value is the arc length of $\vec{r}(t)$ from $t = 0$ to $t = 1$.

h) Find the point on the graph of $\vec{R}(t)$ corresponding to $t = \sqrt{5}$.

i) Find the speed function for $\vec{r}(t)$ and for $\vec{R}(t)$.

Note: One result of reparametrizing a vector function by arc length, gives a speed function which is identically equal to 1. The graph of the reparametrization is identical to the original graph, the only change is the speed on the curve. Thus, in this case the derivative of the reparametrized vector function is identical to the unit tangent vector.

11. Let $\vec{r}(t) = \langle t - 2, \sqrt{t^3} \rangle$.

a) Reparametrize $\vec{r}(t)$ by arc length, using $t = -\frac{4}{9}$ as the basepoint for s.

b) Compute $s^{-1}(0)$.

c) Show that the speed function for the reparametrization is 1.

12. Let $\vec{r}(t) = \left\langle \frac{2}{t^2+1} - 1, \frac{2t}{t^2+1} \right\rangle$. Prove that the curve is the unit circle centered on the origin, by reparametrizing $\vec{r}(t)$ by arc length, using $t = 0$ as the basepoint for s.

Higher Order Derivatives of Functions with Multiple Outputs

Rebekah Muzzi

Derivatives of real-valued multivariable functions can be found using matrices. But when a function has multiple variables and multiple outputs, its derivatives are too complex to be expressed by a single matrix. We can, however, express its derivatives by creating matrices with matrix entries.

The first derivative of a function of the form $f: \mathbb{R}^n \to \mathbb{R}$ is the gradient vector

$$Df = \nabla f = [f_{x_1} \quad f_{x_2} \quad \cdots \quad f_{x_n}].$$

The second derivative is given by the Hessian matrix

$$Hf = \begin{bmatrix} f_{x_1^2} & f_{x_1 x_2} & \cdots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2^2} & & f_{x_1 x_2} \\ \vdots & & \ddots & \vdots \\ f_{x_n x_1} & \cdots & & f_{x_n^2} \end{bmatrix}.$$

For $f \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ where $f_k = f_k(x_1, \ldots, x_n)$, the first derivative, also called the total derivative or the Jacobian matrix, is given by

$$Df = \begin{bmatrix} \nabla f_1 \\ \nabla f_2 \\ \vdots \\ \nabla f_m \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}$$

We can extend this pattern to find the second derivative of functions of the form $f: \mathbb{R}^n \to \mathbb{R}^m$ by listing the Hessian of each component function.

For $f \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ where $f_k = f_k(x_1, \ldots, x_n)$, the second derivative is given by $D^2 f = \begin{bmatrix} Hf_1 \\ Hf_2 \\ \vdots \\ Hf_m \end{bmatrix}.$

Example: Find Df and $D^2 f$ for $f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} xy^3 + z^5 \\ xye^z \\ 3x^3 + y^2 z^4 \end{pmatrix}$

Solution: First, find the gradient vectors of $f_1, f_2,$ and f_3 and combine them to make Df.

$$\nabla f_1 = [y^3 \quad 3xy^2 \quad 5z^4] \quad \nabla f_2 = [ye^z \quad xe^z \quad xye^z] \quad \nabla f_3 = [9x^2 \quad 2yz^4 \quad 4y^2z^3]$$

$$Df = \begin{bmatrix} [y^3 & 3xy^2 & 5z^4] \\ [ye^z & xe^z & xye^z] \\ [9x^2 & 2yz^4 & 4y^2z^3] \end{bmatrix}$$

Then find the Hessian matrices of $f_1, f_2,$ and f_3.

$$Hf_1 = \begin{bmatrix} 0 & 3y^2 & 0 \\ 3y^2 & 6xy & 0 \\ 0 & 0 & 5z^4 \end{bmatrix} \quad Hf_2 = \begin{bmatrix} 0 & e^z & ye^z \\ e^z & 0 & xe^z \\ ye^z & xe^z & xye^z \end{bmatrix} \quad Hf_3 = \begin{bmatrix} 18x & 0 & 0 \\ 0 & 2z^4 & 8yz^3 \\ 0 & 8yz^3 & 12y^2z^2 \end{bmatrix}$$

Finally, combine the Hessians into one matrix, D^2f.

$$D^2f = \begin{bmatrix} \begin{bmatrix} 0 & 3y^2 & 0 \\ 3y^2 & 6xy & 0 \\ 0 & 0 & 5z^4 \end{bmatrix} \\ \begin{bmatrix} 0 & e^z & ye^z \\ e^z & 0 & xe^z \\ ye^z & xe^z & xye^z \end{bmatrix} \\ \begin{bmatrix} 18x & 0 & 0 \\ 0 & 2z^4 & 8yz^3 \\ 0 & 8yz^3 & 12y^2z^2 \end{bmatrix} \end{bmatrix}$$

We can verify that the second derivative works by using it to find approximations. First, we need a formula for approximating function values with second derivatives. The second difference quotient for a function of the form $f: \mathbb{R} \to \mathbb{R}$ is the difference quotient of the difference quotient.

$$D^2(f,x;h) = \frac{D(f,x+h;h) - D(f,x;h)}{h} = \frac{\dfrac{f(x+h+h) - f(x+h)}{h} - \dfrac{f(x+h) - f(x)}{h}}{h}$$

$$= \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}$$

When h is small, the second derivative is given by the limit

$$f''(x) = \lim_{h \to \infty} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}$$

and $f''(x)$ can be approximated by

$$f''(x) \approx \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}.$$

Solving for $f(x + h)$,

$$f(x + h) \approx -\frac{1}{2}h^2 f''(x) + \frac{1}{2}f(x + 2h) + \frac{1}{2}f(x)$$

Since we are working with multivariable functions, we will use the equivalent formula for vectors where the dot product is used for multiplication:

$$f(\vec{p} + \vec{h}) \approx -\frac{1}{2}\vec{h}^2 f''(\vec{p}) + \frac{1}{2}f(\vec{p} + 2\vec{h}) + \frac{1}{2}f(\vec{p})$$

Example: Find an approximation for $f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix}$ where $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} xy^3 + z^5 \\ xye^z \\ 3x^3 + y^2 z^4 \end{pmatrix}$

Solution: Identify $\vec{p} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{h} = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix}$

From the previous example,

$$D^2 f = \begin{bmatrix} \begin{bmatrix} 0 & 3y^2 & 0 \\ 3y^2 & 6xy & 0 \\ 0 & 0 & 5z^4 \end{bmatrix} \\ \begin{bmatrix} 0 & e^z & ye^z \\ e^z & 0 & xe^z \\ ye^z & xe^z & xye^z \end{bmatrix} \\ \begin{bmatrix} 18x & 0 & 0 \\ 0 & 2z^4 & 8yz^3 \\ 0 & 8yz^3 & 12y^2z^2 \end{bmatrix} \end{bmatrix} \qquad D^2 f|_p = \begin{bmatrix} \begin{bmatrix} 0 & 3 & 0 \\ 3 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix} \\ \begin{bmatrix} 0 & e & e \\ e & 0 & e \\ e & e & e \end{bmatrix} \\ \begin{bmatrix} 18 & 0 & 0 \\ 0 & 2 & 8 \\ 0 & 8 & 12 \end{bmatrix} \end{bmatrix}$$

Substituting into the formula,

$$f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix} \approx -\frac{1}{2} \begin{bmatrix} \begin{bmatrix} 0 & 3 & 0 \\ 3 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix} \\ \begin{bmatrix} 0 & e & e \\ e & 0 & e \\ e & e & e \end{bmatrix} \\ \begin{bmatrix} 18 & 0 & 0 \\ 0 & 2 & 8 \\ 0 & 8 & 12 \end{bmatrix} \end{bmatrix} \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix}\begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix} + \frac{1}{2}f\begin{pmatrix} 1.2 \\ 1.2 \\ 1.2 \end{pmatrix} + \frac{1}{2}f\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= -\frac{1}{2} \begin{bmatrix} 1.8 & 0.572 & 0.272 \\ 0.572 & 0.8 & 1.072 \\ 0.272 & 1.072 & 1.972 \end{bmatrix} \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 4.562 \\ 4.781 \\ 8.170 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 2.718 \\ 4 \end{pmatrix}$$

$$= -\frac{1}{2} \begin{pmatrix} 0.2644 \\ 0.2444 \\ 0.3316 \end{pmatrix} + \begin{pmatrix} 2.281 \\ 2.391 \\ 4.085 \end{pmatrix} + \begin{pmatrix} 1 \\ 1.359 \\ 2 \end{pmatrix} = \begin{pmatrix} 3.149 \\ 3.627 \\ 5.919 \end{pmatrix}$$

Check the approximation by directly evaluating $f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix}$

$$f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix} = \begin{pmatrix} 3.075 \\ 3.635 \\ 5.76 \end{pmatrix} \approx \begin{pmatrix} 3.149 \\ 3.627 \\ 5.919 \end{pmatrix}$$

Before we can find third and higher order derivatives of functions with multiple outputs, we need higher order derivatives of multivariable functions with single outputs. Let's observe the pattern so far:

First Derivative:

$$Df = \nabla f = [f_{x_1} \quad f_{x_2} \quad \cdots \quad f_{x_n}]$$

Second Derivative:

$$Hf = \begin{bmatrix} f_{x_1^2} & f_{x_1 x_2} & \cdots & f_{x_1 x_n} \\ f_{x_2 x_1} & f_{x_2^2} & & f_{x_1 x_2} \\ \vdots & & \ddots & \vdots \\ f_{x_n x_1} & \cdots & & f_{x_n^2} \end{bmatrix}$$

We can also write this as

$$Hf = \begin{bmatrix} \dfrac{\partial}{\partial x_1}[f_{x_1} \quad f_{x_2} \quad \cdots \quad f_{x_n}] \\ \dfrac{\partial}{\partial x_2}[f_{x_1} \quad f_{x_2} \quad \cdots \quad f_{x_n}] \\ \vdots \\ \dfrac{\partial}{\partial x_n}[f_{x_1} \quad f_{x_2} \quad \cdots \quad f_{x_n}] \end{bmatrix} = \begin{bmatrix} \dfrac{\partial}{\partial x_1}Df \\ \dfrac{\partial}{\partial x_2}Df \\ \vdots \\ \dfrac{\partial}{\partial x_n}Df \end{bmatrix}.$$

So the next derivative can be found by taking the partial derivatives of the previous derivative with respect to each variable.

The third derivative of a function with a single output is

$$D^3f = \begin{bmatrix} \dfrac{\partial}{\partial x_1}Hf \\[6pt] \dfrac{\partial}{\partial x_2}Hf \\[6pt] \vdots \\[6pt] \dfrac{\partial}{\partial x_n}Hf \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} f_{x_1^3} & f_{x_1x_1x_2} & \cdots & f_{x_1x_1x_n} \\ f_{x_1x_2x_1} & f_{x_1x_2^2} & & f_{x_1x_2x_n} \\ \vdots & & \ddots & \vdots \\ f_{x_1x_nx_1} & & \cdots & f_{x_1x_n^2} \end{bmatrix} \\ \begin{bmatrix} f_{x_2x_1^2} & f_{x_2x_1x_2} & \cdots & f_{x_2x_1x_n} \\ f_{x_2x_2x_1} & f_{x_2^3} & & f_{x_2x_1x_2} \\ \vdots & & \ddots & \vdots \\ f_{x_2x_nx_1} & & \cdots & f_{x_2x_n^2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} f_{x_nx_1^2} & f_{x_nx_1x_2} & \cdots & f_{x_nx_1x_n} \\ f_{x_nx_2x_1} & f_{x_nx_2^2} & & f_{x_nx_1x_2} \\ \vdots & & \ddots & \vdots \\ f_{x_nx_nx_1} & & \cdots & f_{x_n^3} \end{bmatrix} \end{bmatrix}$$

and the fourth derivative is

$$D^4f = \begin{bmatrix} \dfrac{\partial}{\partial x_1}\left[\dfrac{\partial}{\partial x_1}Hf \quad \dfrac{\partial}{\partial x_2}Hf \quad \cdots \quad \dfrac{\partial}{\partial x_n}Hf\right] \\[10pt] \dfrac{\partial}{\partial x_2}\left[\dfrac{\partial}{\partial x_1}Hf \quad \dfrac{\partial}{\partial x_2}Hf \quad \cdots \quad \dfrac{\partial}{\partial x_n}Hf\right] \\[10pt] \vdots \\[10pt] \dfrac{\partial}{\partial x_n}\left[\dfrac{\partial}{\partial x_1}Hf \quad \dfrac{\partial}{\partial x_2}Hf \quad \cdots \quad \dfrac{\partial}{\partial x_n}Hf\right] \end{bmatrix} = \begin{bmatrix} \dfrac{\partial^2(Hf)}{\partial x_1^2} & \dfrac{\partial^2(Hf)}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2(Hf)}{\partial x_1\partial x_n} \\[10pt] \dfrac{\partial^2(Hf)}{\partial x_2\partial x_1} & \dfrac{\partial^2(Hf)}{\partial x_2^2} & & \dfrac{\partial^2(Hf)}{\partial x_2\partial x_n} \\[6pt] \vdots & & \ddots & \vdots \\[6pt] \dfrac{\partial^2(Hf)}{\partial x_n\partial x_1} & \dfrac{\partial^2(Hf)}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2(Hf)}{\partial x_n^2} \end{bmatrix}$$

Since even derivatives will be square matrices like the Hessian, we will notate them in terms of H. If k is even,

$$D^kf = H^{\frac{k}{2}}f$$

Example: $D^4f = H^2f$

So the general formula for the kth derivative is

$$D^kf = \begin{bmatrix} \dfrac{\partial}{\partial x_1}H^{\frac{k-1}{2}}f \\[10pt] \dfrac{\partial}{\partial x_2}H^{\frac{k-1}{2}}f \\[8pt] \vdots \\[8pt] \dfrac{\partial}{\partial x_n}H^{\frac{k-1}{2}}f \end{bmatrix} \text{ if } k \text{ is odd and}$$

$$D^kf = H^{\frac{k}{2}}f = \begin{bmatrix} \dfrac{\partial^2}{\partial x_1^2}H^{\frac{k-2}{2}}f & \dfrac{\partial^2}{\partial x_1\partial x_2}H^{\frac{k-2}{2}}f & \cdots & \dfrac{\partial^2}{\partial x_1\partial x_n}H^{\frac{k-2}{2}}f \\[10pt] \dfrac{\partial^2}{\partial x_2\partial x_1}H^{\frac{k-2}{2}}f & \dfrac{\partial^2}{\partial x_2^2}H^{\frac{k-2}{2}}f & & \dfrac{\partial^2}{\partial x_2\partial x_n}H^{\frac{k-2}{2}}f \\[6pt] \vdots & & \ddots & \vdots \\[6pt] \dfrac{\partial^2}{\partial x_n\partial x_1}H^{\frac{k-2}{2}}f & \dfrac{\partial^2}{\partial x_n\partial x_2}H^{\frac{k-2}{2}}f & \cdots & \dfrac{\partial^2}{\partial x_n^2}H^{\frac{k-2}{2}}f \end{bmatrix} \text{ if } k \text{ is even.}$$

Now we can find higher order derivatives of functions with multiple outputs. The kth derivative of a function of the form $f: \mathbb{R}^n \to \mathbb{R}^m$ is the column vector where the jth entry is the kth derivative of the jth component function.

Example: Find $D^3 f$ of $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x^3 y^3 \\ x^4 + y^4 \end{pmatrix}$

Solution: First, find the partial derivatives of f.

$$\frac{\partial f_1}{\partial x} = 15x^2 y^3 \qquad \frac{\partial f_1}{\partial y} = 15x^3 y^2$$

$$\frac{\partial f_2}{\partial x} = 4x^3 \qquad \frac{\partial f_2}{\partial x} = 4y^3$$

Then find the Hessian of each component function.

$$Hf_1 = \begin{bmatrix} 30xy^3 & 45x^2 y^2 \\ 45x^2 y^2 & 30x^3 y \end{bmatrix} \qquad Hf_2 = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$$

Then take the partial derivatives with respect to x and y of each Hessian matrix to find $D^3 f$.

$$D^3 f = \begin{bmatrix} \begin{bmatrix} 30y^3 & 90xy^2 \\ 90xy^2 & 60x^2 y \end{bmatrix} & \begin{bmatrix} 90xy^2 & 90x^2 y \\ 90x^2 y & 30x^3 \end{bmatrix} \\ \begin{bmatrix} 24x & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 24y \end{bmatrix} \end{bmatrix}$$

We can verify the third derivative with the following approximation formula found by taking the difference quotient of the second difference quotient:

$$f(\vec{p} + \vec{h}) \approx \frac{1}{3}\vec{h}^3 f'''(p) - \frac{1}{3}f(\vec{p} + 3\vec{h}) + f(\vec{p} + 2\vec{h}) + \frac{1}{3}f(\vec{p})$$

Example: Find an approximation for $f\begin{pmatrix} 1.1 \\ 1.1 \end{pmatrix}$ where $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x^3 y^3 \\ x^4 + y^4 \end{pmatrix}$

Solution: Identify $\vec{p} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{h} = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$

From the previous example,

$$D^3 f = \begin{bmatrix} \begin{bmatrix} 30y^3 & 90xy^2 \\ 90xy^2 & 60x^2 y \end{bmatrix} & \begin{bmatrix} 90xy^2 & 90x^2 y \\ 90x^2 y & 30x^3 \end{bmatrix} \\ \begin{bmatrix} 24x & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 24y \end{bmatrix} \end{bmatrix} \qquad D^3 f|_p = \begin{bmatrix} \begin{bmatrix} 30 & 90 \\ 90 & 60 \end{bmatrix} & \begin{bmatrix} 90 & 90 \\ 90 & 30 \end{bmatrix} \\ \begin{bmatrix} 24 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 24 \end{bmatrix} \end{bmatrix}$$

Substitute into the approximation formula

$$f(\vec{p} + \vec{h}) \approx \frac{1}{3}\left[\begin{bmatrix} \begin{bmatrix} 30 & 90 \\ 90 & 60 \end{bmatrix} & \begin{bmatrix} 90 & 90 \\ 90 & 30 \end{bmatrix} \\ \begin{bmatrix} 24 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 24 \end{bmatrix} \end{bmatrix}\right]\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} - \frac{1}{3}f(1.3) + f(1.2) + \frac{1}{3}f(1)$$

$$= \frac{1}{3}\left[\begin{bmatrix} \begin{bmatrix} 12 & 18 \\ 18 & 9 \end{bmatrix} \\ \begin{bmatrix} 2.4 & 0 \\ 0 & 2.4 \end{bmatrix} \end{bmatrix}\right]\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 24.134 \\ 5.712 \end{pmatrix} + \begin{pmatrix} 14.93 \\ 4.147 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

$$= \frac{1}{3}\begin{bmatrix} 1.44 & 1.8 \\ 1.8 & 1.14 \end{bmatrix}\begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} - \begin{pmatrix} 8.045 \\ 1.904 \end{pmatrix} + \begin{pmatrix} 14.93 \\ 4.147 \end{pmatrix} + \begin{pmatrix} 1.667 \\ 0.667 \end{pmatrix}$$

$$= \begin{pmatrix} 0.108 \\ 0.098 \end{pmatrix} - \begin{pmatrix} 8.045 \\ 1.904 \end{pmatrix} + \begin{pmatrix} 14.93 \\ 4.147 \end{pmatrix} + \begin{pmatrix} 1.667 \\ 0.667 \end{pmatrix} = \begin{pmatrix} 8.66 \\ 3.008 \end{pmatrix}$$

Check by evaluating $f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix}$

$$f\begin{pmatrix} 1.1 \\ 1.1 \\ 1.1 \end{pmatrix} = \begin{pmatrix} 8.858 \\ 2.928 \end{pmatrix} \approx \begin{pmatrix} 8.66 \\ 3.008 \end{pmatrix}$$

The other derivatives can be verified using similar formulas.

Functions with multiple outputs are impossible for our human minds to visualize, but by using derivatives with matrices embedded within matrices, we can discover useful information about functions in higher dimensions. We explored what the derivatives look like and how to use them to find approximations. These derivatives could also be useful for finding other information about functions such as their maximums, minimums, or curvature. In spite of the limitations of the human imagination, math makes exploring higher dimensions possible.

Exoplanet 51 Pegasi-b, Determination of Orbital Parameter in an Astronomy Laboratory

James Mattiello, Ph.D.

As an Adjunct Instructor at SC4, I have the opportunity to teach the introductory Physics and Astronomy classes. To stimulate the critical thinking skill of my astronomy students, I like to ask a few basic questions to test the astronomical knowledge. First question, "What is the name of the nearest star?". Most of the time the reply is Proxima or Alpha Centauri, which is wrong, the nearest star is our own Sun, *Sol*. I also like to ask, "How many planets do we know about?". This stimulate a debate, is *Pluto* included in the count, should the other newly discovered Kuiper Belt Objects (dwarf-planets) be included (*Haumea, Makemake,* and *Eris*). The answer usually varies between 8 to 14 planets. The students are surprise when I tell them that there are off by a few thousand planets. I didn't restricted the planet count to just our solar system. The actual number is in around 3322, as of 2016. These extra solar system planets, *exoplanets*, 9 from our own solar system, including Pluto, and some 3300 exoplanets around other stars. All these exoplanets have been confirmed via astronomical observation and are located within some 30,000 light years distance from Earth. You can check the most current number by visiting the exoplanets official websites, http://exoplanets.org or Hanno Rein's exoplanet app, http://exoplanetapp.com.

In my astronomy class we learn about the four basic methods used to detection exoplanets; radial velocity measurements, transit light curves, direct imaging, and gravitational micro-lensing, (Perryman, 2011). The motion of a planet in orbit around a host star causes this star to undergo a reflex motion about the star-planet center of mass. Additional planets also produce additional reflex motion about the center of mass. This motion results in the periodic perturbation of observable properties, radial velocity, angular position in the sky, and in the arrival time of periodic reference signal from the host star.

Most of the exoplanets discovered to date have been detected using the radial velocity method. This method uses precise spectroscopic measurements of the periodic Doppler shift of a reference spectral lines of the host star. The line of sight, radial part, of this Doppler shift motion occurs because the host star is moving around the center of mass. An away from the Earth motion, produces a Doppler redshift (shift toward longer wavelengths) of the spectral line. A toward the Earth motion, produces a Doppler blueshift (shift to shorter wavelengths) of the spectral line. The Doppler shift equation used by astronomers to calculate the radial velocity is;

$$V = \left(\frac{\lambda_{obs} - \lambda_{ref}}{\lambda_{ref}} \right) c$$

eq. 1

where (λ_{ref}) is the reference wavelength, (λ_{obs}) is the observed wavelength from the host star, and c is the speed of light (c = *3.0×10^8 m/s*). Note that a Doppler blueshift produces a negative radial velocity, while a Doppler redshift produces a positive radial velocity.

In the astronomy laboratory, the students use published data from radial velocity measurement of star 51 Pegasi, and its exoplanet 51 Pegasi-b (Marcy, GW, Butler RP, 1997). This star is in the constellation Pegasus and is similar in mass, temperature and radius to our own Sun. The date and time for these observation are in a Julian date format. The Julian date is used by astronomers as a way of expressing the date and time as a single number, counting the days (and fraction of days) since January 1st, 4713 BC. These observation started on the Julian date 2,450,000 which was from October 11, 1995 and through August 9, 1996.

The students laboratory assignment is to fit a non-linear function to the radial velocity data. From which they can determine the orbital period, the planet's time to complete one orbit around the host star, and the other parameters for this exoplanet, its mass and semi-major axis (distance from the host star). From a preliminary examination of the 51 Pegasi's data, it is apparent that it is sinusoidal in nature, with an orbital period of approximately *4 days* and a radial velocity of about ± *55 m/s*. If you want to fit this data using an Microsoft excel or Apple's numbers spreadsheet program, a problem arises that these spreadsheets don't fit a sinusoidal function to the data. To solve this issues, the computing software Mathematica˙ is used to process the data, the Mathematica code is shown in Appendix C at the end of this article.

The students begin by entering the Julian date (time) data, as a row array, this array is stored as a variable. The radial velocity data are entered as another row array, and are stored as a second variable. The dimension for the two arrays are checked to making sure that both list have the same number of data points. The two arrays are combined and put into a (x, y) column matrix. Where x is the *date (time)* and y is the *radial velocity*. A non-linear sinusoidal function of the following form is used to fit the data;

$$y = A\sin\left(\left(\tfrac{2\pi}{B}\right)x - C\right) + D$$

eq. 2

Where the amplitude *A* is the radial velocity (in *m/s*), *B* is the orbital period (in *days*), *C* is the phase and *D* the y-intercept. The first attempt to fit this function to the data results in a poor outcome, see Figure 1.
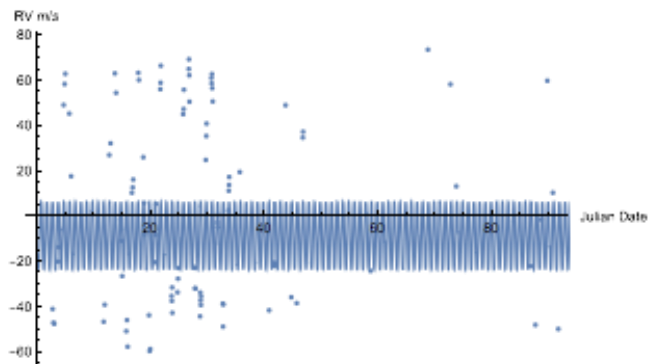

Figure 1.
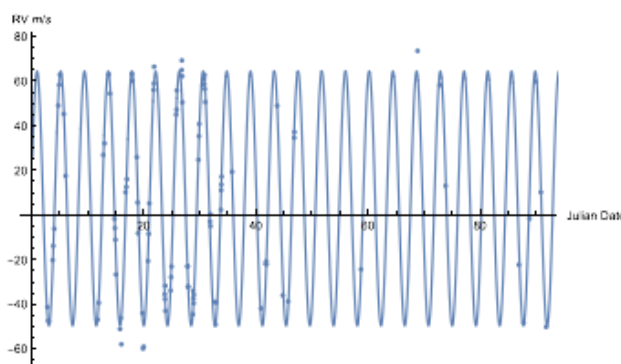

Figure 2.

Remember that this data is sinusoidal in nature, with a period of approximately 4 days and a radial velocity amplitude of about ± 55 m/s. For a second attempt, the non-linear fitting algorithm is given some assistance, using a guess of A = *55 m/s* and B = *4 days* as starting values. This second attempt results in a better non-linear fit to the data as shown in Figure 2.

This non-linear fit to the sinusoidal function (equation 2) results in a set of the parameters, the orbital period (B = *4.230 days*), and a radial velocity of (A = ± *56.978 m/s*) which the students are prepared to use to determine the other exoplanet properties. The orbital period for exoplanet 51 Pegasi-b, *4.230 days* is converted into years, *0.01158 years*. Using Kepler's 3rd law, *the orbital period (in years) squared is equal to the semi-major axis (in AU) cubed*, one can determine the semi-major axis. Remember the Earth is 1 *Astronomical Units, 1 AU,* from our Sun with an orbital period of *1 year*. Using the converted orbital period value the semi-major axis is calculated to be *0.0512 AU*. The mass of the exoplanet is calculated using a modified Kepler's 3rd laws equation, which uses the Doppler shift radial velocity of the host star and the exoplanet's orbital period, the equation is derived in Appendix A and B. From this equation, the mass of 51 Pegasi-b is determine to be *0.45 M_J.* It is calculated in Jupiter masses to compare this exoplanet's properties to values observed for our own solar system.

The students have derived the three basic parameters of this exoplanet 51 Pegasi-b its, mass (*0.45 M_J*), orbital period (*4.230 days*), and semi-major axis (*0.0512 AU*). In keeping with Cindy Wade's philosophy

published in the first volume of this journal, "And Check Your Work". The students are reminded to compare their calculated values with published values. The host star properties are determined from stellar evolution models and other astronomical observations and are summarized in Table 1. Also shown in Table 1 are published properties of the exoplanet 51 Pegasi-b. From the values listed in Table 1, the experimental error for the calculated orbital period and semi-major axis is less than 1%, and the experimental error for the mass was approximately 2%. The students are very pleased with their results.

Remember that 51 Pegasi is similar is size to our own Sun. So one might expect a similar solar system, and by u sing our solar system and its giant planets as a reference model. The students are asked to reflex on a few questions to complete this astronomy laboratory;

1. Considering the mass of this planet as well as its distance from the host star, why was the discovery of exoplanet 51 Pegasi-b such a surprise?

2. Comment on the environment of this planet?

3. Is it possible for this exoplanet to be hospitable to life?

As a reference, Mercury has an orbital period of *0.241 years* (*88 days*) and is *0.387 AU* from the Sun, while Jupiter has an orbital period of *11.863 years* (about *4333 days*) and is *5.203 AU*. It should be noted that this radial velocity method has a bias in the detection of massive planets with small orbits, they are commonly referred to as *"Hot Jupiters"*. Compared to our solar system, where our giant planets are at large distances (> *5 AU*) from our Sun. The location of these Hot Jupiters gives rise to an astronomical mystery;

• "Why do these Hot Jupiters exist in their small orbits?"

• "Do these exoplanets migration towards their host star or did they form at these small orbits?"

Recently the international Astronomical Union has allowed the public the opportunity to name some of these star systems and their exoplanets. (Montmerle T, Benventuti P, 2015) The host star 51 Pegasi has been named, Helvetios, and its exoplanet 51 Pegasi-b was named, Dimidium. (Beatty JK, 2016).

References

Rein H, "Exoplanet app", http://exoplanetapp.com, 2016.

Perryman M, *The Exoplanet Handbook*, Cambridge University Press, 2011.

Marcy, GW, Butler RP, et.al., "The Planet Around 51 Pegasi", *The Astrophysical Journal*, 481: 926 - 935, 1997 June 1.

Wade C, "And Check your Work", Continuum, vol. 1, num. 1: 41 - 42, 2016.

Montmerle T, Benventuti P, et. al., "Public Naming of Planets and Planetary Satellites", Transactions IAU, Vol. XXIXA, Proc. XXVIII IAU General Assembly, August 2012..

Beatty JK, Sky & Telescope, April 2016, p 14.

Table 1
reference http://exoplanets.org/detail/51_Peg_b

51 Pegasi properties		51 Pegasi-b properties	
Mass of Star	$1.054 \pm 0.03\ M_\odot$	Mass of Exoplanet	$0.461 \pm 0.016\ M_J$
Radius of Star	$1.025 \pm 0.036\ R_\odot$	Semi-major Axis	$0.0521 \pm 0.0008\ AU$
[Fe/H]	0.200 ± 0.03	Orbital Period	$4.2311 \pm 0.0005\ days$
Stellar Spectral Type	G2 IV		
Temperature	5787 ± 44 K		
Radial Velocity	$55.9 \pm 0.8\ m/s$		
Distance to Star	$15.61 \pm 0.01\ pc$ ~ 50.9 *light years*		
$M_\odot$ and $R_\odot$ represents a solar mass and solar radius.		M_J represent a Jupiter mass.	

Appendix A.

If we assume that the exoplanet system is observed in the plane of the exoplanet's orbit, we can then use the following center of mass equation to calculate the exoplanet's mass, $M_{star}V_{star} = M_{planet}V_{planet}$. Where the parameter are the host star's mass and Doppler shift radial velocities and the exoplanet's mass and orbital speed. If the system is not view in the plane of the exoplanet but at some inclined angle, it is not possible to measure the non-radial part of the radial velocity. It is only possible to obtain a lower limit on the mass of the exoplanet and so its mass might actually be larger, $M_{planet,real} \geq M_{planet,real}\sin(i) = M_{planet}$, where i is the angle of inclination.

To derive a form of Kepler's 3rd law for determining the mass of a planet. We start by deriving an equation of the star's mass in terms of the orbital speed and semi-major axis of the exoplanet;

$$F_{Gravity} = G\frac{M_{star}M_{planet}}{R^2} = \frac{M_{planet}V^2_{orbital\ speed}}{R} = F_{circular\ motion}$$

$$M_{star} = \frac{RV^2_{orbital\ speed}}{G}$$

where R is the semi-major axis (in *m*), $V_{orbital\ speed}$ is the exoplanet's orbital speed (in *m/s*) and G is the universal gravitational constant (6.67×10^{-11} Nm^2/kg^2). Using this value of the star's mass (M_{star}), and the center of mass equation, we derive the equation for calculating the mass of the exoplanet, knowing the Doppler shift radial velocity of the host star and the orbital period of the exoplanet.

$$M_{planet}V_{orbital\ speed} = M_{star}V_{star\ doppler\ shift}$$

$$M_{planet}V_{orbital\ speed} = \left(\frac{RV_{orbital\ speed}^2}{G}\right)V_{star\ doppler\ shift}$$

$$M_{planet} = \frac{R}{G}V_{orbital\ speed}V_{star\ doppler\ shift} \qquad\qquad with\ V_{orbital\ speed} = \frac{2\pi R}{T}$$

$$M_{planet} = \frac{2\pi}{G}\frac{R^2}{T}V_{star\ doppler\ shift} \qquad\qquad Kepler's\ 3rd\ Law,\ R^3 = T^2$$

$$M_{planet} = \frac{2\pi}{G}\frac{(T^{2/3})^2}{T}V_{star\ doppler\ shift} = \frac{2\pi}{G}T^{1/3}V_{star\ doppler\ shift} \qquad\qquad eq.\ 5$$

The Doppler shift radial velocity for our Sun due to the effect of Jupiter is ($V_{Sun} = 12.45\ m/s$), and is calculated in Appendix B. We want to compare the exoplanet mass to the mass of Jupiter, so our final equation becomes;

$$\frac{M_{planet}}{M_{Jupiter}} = \frac{\frac{2\pi}{G}T_{planet}^{1/3}V_{star\ doppler\ shift}}{\frac{2\pi}{G}T_{Jupiter}^{1/3}V_{Sun\ doppler\ shift}}$$

$$M_{planet} = \left(\frac{T_{planet}}{T_{Jupiter}}\right)^{1/3}\left(\frac{V_{star\ doppler\ shift}}{V_{Sun\ doppler\ shift}}\right)M_{Jupiter} = \left(\frac{0.01158\ yr}{11.863\ yr}\right)^{1/3}\left(\frac{56.98}{12.45}\right)M_{Jupiter} = 0.45\ M_{Jupiter} \qquad eq.\ 6$$

Appendix B.

To calculating the Doppler shift radial velocity of our Sun due to Jupiter, we use the center of mass equation;

$$M_{Sun}V_{Sun} = M_{Jupiter}V_{Jupiter}$$

$$R_{Jupiter} = 5.203\ AU\frac{1.496\times10^{11}m}{1\ AU} = 7.784\times10^{11}m$$

$$V_{Sun} = \frac{M_{Jupiter}}{M_{Sun}}\left(\frac{2\pi R}{T_{Jupiter}}\right)$$

$$T_{Jupiter} = 11.863\ yr\frac{3.16\times10^7 s}{1\ yr} = 3.749\times10^8 s$$

$$V_{Sun} = \frac{1.898\times10^{27}kg}{1.989\times10^{30}kg}\left(\frac{2\pi(7.784\times10^{11}m)}{(3.749\times10^8 s)}\right) = 12.449\ \frac{m}{s}$$

Sophie Germain

Marie-Sophie Germain was born on April 1st in Paris in 1776. (Louis XVI and his wife Marie Antoinette had assumed power ten months earlier.) She was the middle of three sisters. She was fascinated by lectures delivered by Lagrange at the École Polytechnique, and wrote brilliant letters to him signed "M. (monsieur) Le Blanc." Women were not allowed to attend the École at this time, prompting biographer Dora Musielak to ask, "How did Germain learn mathematics on her own before sending her analysis to Lagrange?" (Musielak, Prime Mystery: The Life and Mathematics of Sophie German, © 2015, p. viii.)

> That was a time when France was perhaps the most powerful country in Europe, and its cultural influence was such that nobles, monarchs, and the educated people in many other countries often spoke elegant French instead of their native languages. France was also among the most scientifically enlightened nations in the world. The Royal Academy of Sciences in Paris had among its members some of the most eminent and influential mathematicians in history such as Lagrange, Laplace, and Legendre (Luh-ZHON-dra.) (Musielak, p. 1)

Sophie Germain lived 500 meters from La Conciergerie, the prison where Marie Antoinette and many others awaited their turn with the guillotine.(Note: the Paris Academy of Sciences devised the metric system around this time!) Her father was a politician, possibly a revolutionary, and her childhood would have been filled with overheard politics relating the French Revolution. She hid in her father's library and read Archimedes.

Lagrange's printed lecture notes became available in 1798 (Musielak p. 33). He wrote about calculus and his idea of how to "unite calculus and algebra" and free differential calculus from "the difficulties that occur….that discourage most of those who undertake its study." (Musielak p. 28.) Around this time Germain wrote letters to Lagrange commenting on these lectures.

In an 18th century example of "mansplaining," Professor J.A.J Cousin sent Germain an algebra text he had written and offered to help her with her studies – Sophie at 21 was already far beyond this text. Another Parisian scientist apparently insulted her; he would later write

> Il était difficile, Mademoiselle, de me faire sentir plus que vous ne l'avez fait hier, l'indiscretion de ma visite et l'improbation de mes hommages, mais il m'était difficile de le prévoir. (Musielak, p. 34.)

> (It would be difficult, Mademoiselle, for anyone to make me feel as you did yesterday, for the indiscretion of my visit and the disapproval of my respects to you, but it was difficult for me to predict it.)

The great mathematician Gauss wrote this in a letter to Germain:

> But when a person of the sex, which according to our customs and prejudices, must encounter infinitely more difficulties than men to familiarize herself with these thorny researches, succeeds nevertheless in surmounting these obstacles and penetrating the most obscure parts of them, then without doubt she must have the noblest courage, quite extraordinary talents and a superior genius. (Lynn Osen, Women in Mathematics, quoting from E.T. Bell, quoting Gauss.)

Sophie Germain produced results in number theory, particularly regarding Fermat's Last Theorem, and Sophie Germain primes are named for her.

Definition: A prime number p is a Sophie Germain prime if $2p + 1$ is also prime.

The largest Sophie Germain prime known has 388,342 digits. From Wolfram Mathworld:

"Around 1825, Sophie Germain proved that the first case of Fermat's last theorem is true for such primes, i.e., if p is a Sophie Germain prime, then there do not exist integers x,y, and z different from 0 and none a multiple of p such that

$$x^p + y^p = z^p$$

Sophie Germain's Identity states that $x^4 + 4y^4 = ((x + y)^2 + y^2)((x - y)^2 + y^2)$

Germain died of breast cancer in 1831.

Questions:

1.   Where did young Sophie hide during the French Revolution?

2.   What great French mathematician received letters from Germain on his lectures?

3.   Did that mathematician consider differential calculus difficult for most students?

4.   What scientific organization whose members included Laplace and Lagrange existed in Paris at this time?

5.   How did Professor Cousin condescend to Germain?

6.   How does the Declaration of Independence help us remember something about Sophie Germain?

7.   Did Gauss admire her?

8.   Is 11 a Sophie Germain prime? 13? 17? Verify your answers.

9.   Verify Sophie Germain's Identity directly using algebra.